Studying the structural features of peptide dendrimers using a combined computational and experimental approach

Luís Carlos Santos Filipe

supervised by
Dr. António M. BAPTISTA
Dr. Tamis DARBRE
Dr. Miguel MACHUQUEIRO

Oeiras, 2016

Para os meus avós maternos

Acknowledgments

I would like to start this thesis by acknowledging the contributions made to it by Miguel Machuqueiro, Tamis Darbre and António Baptista. Their teachings, guidance and support over the years were critical to the sound conclusion of the work presented here and to my growth as a "scientist". It is because of them that I could not, in conscience, write this dissertation entirely in the first-person singular and felt compelled to write it in the first-person plural.

A very special word of thanks is ought to António, for welcoming me into is lab and for all his reviewing efforts over the last months.

In the course of this project I have, fortunately, come to know a number of persons who directly or indirectly contribute to it. Probably, they are unaware of how their companionship and every day good humour helped me to stay motivated and maintain some degree of mental sanity. To properly thank them, I am going to break the usual protocol of an acknowledgements section and include some pictures where we are together (that, and because they truly are a lot to mention by name).

I must, however, single out two friends: João Damas and Pedro Magalhães. Without their IT skills and the willingness to use them, this work would have taken twice as long to do.

I also want to acknowledge the two academic institutions that have hosted me: the Instituto de Tecnologia Química e Biológica and the University of Bern. The persons in the academic and scientific staffs of both institutions really ensured the adequate environment to carry out my research. I also acknowledge Fundação para a Ciência e Tecnologia for financial support through grant SFRH/BD/76085/2011.

None of this would have been possible without the support, incen-

The 1st Portuguese Biomolecular Modelling meeting, Tavira, 2011

The Reymond group, Bern, 2013

tives and love of my parents and my brother.

Thanks to Rita and little Vasco, for continuously reminding me what, at the end of the day, is truly important. She, the amazing spice in my every day life, and him, the chaotic rascal whose smile immediately puts everything into perspective. They make me try to be the man I aspire to be.

Abstract

Dendrimers are a family of highly branched synthetic compounds that share a common layout where wedges emerge radially from a core by means of a regular branching pattern. Topologically, dendrimers are characterized by three distinct regions: core, branches, and periphery, with the latter being typically composed of functionalized end-groups.

Peptide dendrimers are dendrimers composed of amino acids. These peptidic tree-like molecules have attracted considerable interest due to their inherent multivalency and the possibility of constructing molecules with a variable number and type of amino acids within the dendrimer branches and core. This topologic versatility, along with the possibility of grafting different functional groups to the dendrimers end-groups, has sparked the development of several applications ranging from catalytic peptide dendrimers and metalloprotein models to antimicrobial and drug delivery agents.

Although the synthesis of these systems is becoming increasingly straightforward, two main issues remain mostly unaddressed and hinder their widespread application: (1) little is known about these molecules structure in solution and the influence played by each of its constituents on the overall fold, and (2) the selection of which amino acids to use for specific functions is ruled by chemical reasoning and trial and error screening of large combinatorial libraries.

There are at least two kinds of preferential structural behaviors exhibited by these molecules, which acquire either compact or noncompact shapes. However, the key structural determinants of such behaviors remained, until now, unstudied.

In the first part of this work, we have conducted a comprehensive

investigation of the structural determinants of peptide dendrimers by employing long molecular dynamics simulations to characterize an extended set of third generation dendrimers. In particular, we investigated the structural role played by branching and the replacement of neutral residues by negatively charged ones. Our results support the idea that a joint effect of electrostatics and hydrogen bonds is the defining factor for these molecules structural behavior in solution.

In the second part of this work, we have combined experimental and computational techniques to study the structure of positively charged peptide dendrimers. Third-generation dendrimers containing combinations of positive/neutral amino acid residues in the different dendrimer generations were synthesized and their compactness evaluated using diffusion NMR. Molecular dynamics simulations were performed to obtain a comprehensive description of the molecular–level phenomena substantiating the structural differences observed. Comparison of the results presented with the first part of the work reveals a striking charge-dependent tendency in these systems, where the simple number and placement of charged amino acids in the sequence allows an extensive control over the exhibited structural features.

pH effects and electrostatic interactions are known to be crucial players to explain the conformational and functional behaviors observed in these systems. In the third part of this work, we present the results of constant-pH molecular dynamics simulations performed at several pH values for four peptide dendrimers of different generations (from one to four) composed of the same type of aminoacids: histidines, serines and diaminopropionic acid. These dendrimers are known to catalyze the hydrolysis of pyrene sulfonate esters. Constant-pH MD simulations in the presence of substrate molecules at the optimum pH for catalysis are also reported.

The results show that first and second generation dendrimers are almost structurally unresponsive to pH variations. For third and fourth generation dendrimers, pH plays a structuring role, with markedly different behaviors being observed when passing from acidic to neutral pH. Protonation–conformation coupling effects influence several intra-

molecular interactions, which, in turn, modulate the shape and structure at the different pH values. The atypical and highly pH-dependent protonation profiles of some histidine residues are also investigated.

The interactions between dendrimers and substrates restrict the conformational space available to the dendrimers and enforces conformational homogeneity. This structuring effect is a consequence of the dendrimer–substrate interactions which occur through stabilizing hydrogen bonds and ion pairs between the substrates sulfonate groups and the dendrimers residues.

In the last chapter of this thesis we discuss the key findings of our studies and their broad implications. In general, the results presented constitute valuable fundamental knowledge allowing a deeper, atomic-level understanding of peptide dendrimers structural features.

Resumo

Os dendrímeros são uma família de compostos sintéticos altamente ramificados, caracterizados por um padrão regular de crescimento a partir de um núcleo comum. Topologicamente os dendrímeros são caracterizados por três regiões distintas: um núcleo, as diferentes ramificações e uma periferia normalmente composta por grupos terminais funcionalizados.

Os dendrímeros peptídicos são dendrímeros constituídos por aminoácidos. Estas moléculas peptídicas e ramificadas têm atraído uma considerável atenção devido à sua multivalência e à possibilidade de sintetizar moléculas com inúmeros e variados aminoácidos nas diferentes ramificações. Tal versatilidade topológica, juntamente com a possibilidade de enxertar diferentes grupos funcionais nos terminais de um dendrímero, tem suscitado interesse nestas moléculas e conduziu já a diversas aplicações. Os dendrímeros peptídicos desempenham nessas aplicações papéis muito variados, desde modelos catalíticos e miméticos de metalo-proteínas, até agentes antimicrobianos ou transportadores de fármacos.

Apesar dos consideráveis avanços na síntese destes sistemas, dois tópicos importantes permanecem inexplorados, dificultando a aplicação e aceitação generalizada destes sistemas. Em primeiro lugar, existe pouca informação acerca da estrutura destas moléculas em solução e do papel desempenhado por cada um dos seus constituintes na aquisição dessa mesma estrutura. Em segundo lugar, a selecção dos aminoácidos a utilizar para obter comportamentos funcionais específicos é baseada essencialmente na experiência adquirida, na intuição química dos interessados e, quando possível, em processos de *screening* de livrarias

combinatórias contendo várias sequências peptídicas.

Em termos estruturais, são conhecidos dois tipos de comportamento preferenciais nestas moléculas, que podem adoptar formas compactas ou abertas. Contudo, os determinantes conformacionais preponderantes na adopção de um comportamento em detrimento de outro permaneciam, até agora, por estudar.

Na primeira parte deste trabalho realizámos uma investigação abrangente dos determinantes conformacionais dos dendrímeros peptídicos. Para tal, realizámos longas simulações de dinâmica molecular por forma a caracterizar um extenso conjunto de dendrímeros de terceira geração. Investigámos especificamente o papel desempenhado pelos resíduos de bifurcação (variando o comprimento da sua cadeia lateral) e pela quantidade e distribuição de carga negativa (substituindo glutaminas por ácidos glutâmicos). Os resultados obtidos suportam a ideia de que as interacções electrostáticas e as pontes de hidrogénio são factores determinantes no comportamento destas moléculas em solução.

Na segunda parte deste projecto combinámos técnicas experimentais e computacionais para estudar a estrutura de dendrímeros peptídicos carregados positivamente. Dendrímeros de terceira geração com diferentes combinações de aminoácidos positivos/neutros nas diferentes gerações foram sintetizados e a sua compactação foi avaliada através de RMN de difusão. Posteriormente, foram realizadas simulações de dinâmica molecular para obter uma descrição dos detalhes moleculares que ocasionam as diferenças observadas. Juntamente com os resultados obtidos na primeira parte do trabalho, estes resultados mostram que a simples variação do número e posicionamento de resíduos carregados permite controlar características estruturais.

É sabido que os efeitos do pH e as interacções electrostáticas são cruciais para explicar o comportamento conformacional e funcional destes sistemas. Na terceira parte deste trabalho, apresentamos os resultados de simulações de dinâmica molecular a pH-constante, realizadas a vários valores de pH, para quatro dendrímeros peptídicos de diferentes gerações (da primeira à quarta). Estes dendrímeros são compostos pelo mesmo tipo de aminoácidos: histidina, serina e ácido diaminopropió-

nico.

Os resultados destas simulações demostraram que, no que respeita a variações estruturais, os dendrímeros de primeira e segunda geração são praticamente insensíveis a alterações de pH. Nos dendrímeros de terceira e quarta geração o pH desempenha um papel estruturante e observam-se comportamentos diferentes ao passar de pH ácidico para neutro. Os efeitos resultantes do acoplamento entre protonações e conformações influenciam diversas interacções intramoleculares, que por sua vez condicionam a estrutura destes dendrímeros a diferentes valores de pH. Os perfis de protonação atípicos e altamente dependentes do pH de alguns resíduos de histidina foram também investigados.

Dados experimentais demonstram que estes dendrímeros catalisam a hidrólise de ésteres de pireno sulfonados. Como tal, foram também realizadas simulações de dinâmica molecular a pH-constante dos dendrímeros na presença dos respectivos substratos.

As interacções entre dendrímeros e substratos restringem o espaço conformacional acessível aos dendrímeros e impõem uma acentuada homogeneidade conformacional. Este efeito estruturante é uma consequência das interacções entre os dendrímeros e os substratos, interacções essas que ocorrem essencialmente através de pares iónicos e pontes de hidrogénio entre os grupos sulfonato dos substratos e os resíduos dos dendrímeros.

No último capítulo desta tese são discutidas as principais conclusões dos nossos estudos bem como as suas implicações. Globalmente, o trabalho aqui apresentado contribui para uma maior compreensão das características estruturais dos dendrímeros peptídicos a nível atómico, fornecendo alguns princípios básicos que poderão auxiliar o desenho destas moléculas.

Contents

List of Figures

List of Tables

List of Abbreviations

Amino acids

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Ornithine	Orn	
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
2,3-diaminopropanoic acid	Dap	
2,4-diaminobutanoic acid	Dab	
4-aminomethyl(benzoic) acid	Amb	

Other abbreviations and constants

CE	continuum electrostatics
COM	center of mass
CpHMD	constant-pH molecular dynamics
FF	force field
GRF	generalized reaction field
MC	Monte Carlo
MD	molecular dynamics
MM	molecular mechanics
NMR	nuclear magnetic resonance
PB	Poisson–Boltzmann
PAMAM	poly(amidoamine) dendrimers
PPI	poly(propylene imine) dendrimers
PME	particle mesh ewald
PPL	poly-L-lysine dendrimers
rms	root mean squared (deviation, fluctuation, etc.)
SPPS	solid phase peptide synthesis

Latin and Greek symbols

D	diffusion coefficient
h	Planck constant
I	ionic strength
k_B	Boltzmann constant
K	kinetic energy
m_i	mass of particle i
N	number of particles in a system
$\mathrm{p}K_{\mathrm{eff},i}$	effective $\mathrm{p}K_\mathrm{a}$ of a titrable site i, i.e., the pH-dependent $\mathrm{p}K_\mathrm{a}$-like quantity expressing its protonation equilibrium
$\mathrm{p}K_{\mathrm{half},i}$	pH value at which a titrable site i is half-protonated
$\mathrm{p}K_{\mathrm{int},i}$	intrinsic $\mathrm{p}K_\mathrm{a}$ of a titrable site i, i.e., its $\mathrm{p}K_\mathrm{a}$ when all other sites are neutral
$\mathrm{p}K_{\mathrm{mod},i}$	$\mathrm{p}K_\mathrm{a}$ of the model compound corresponding to a titrable site i
$\mathcal{H}$	Hamiltonian function
$\mathbf{p}_i$	momentum of particle i
P	pressure
$\mathbf{r}_i$	coordinates of particle i
R	ideal gas constant
R_g	gyration radius
R_h	hydrodynamic radius
t	time
T	absolute temperature
U	potential energy
V	volume
W	electrostatic energy
δ	asphericity
$\varepsilon(\mathbf{r})$	general dielectric constant
μ	chemical potential
$\xi(\mathbf{r})$	counter-ion charge density
$\rho(\mathbf{r})$	general charge density
$\phi(\mathbf{r})$	general electrostatic potential

Chapter 1

Introduction

Contents

1.1 Dendrimers: concepts, definitions and examples

Nowadays, traditional synthetic polymers are usually classified into four major groups according to their architecture (Figure 1.1). Dendrimers are an important subclass of the fourth major polymer architecture category: dendritic polymers [1–5].

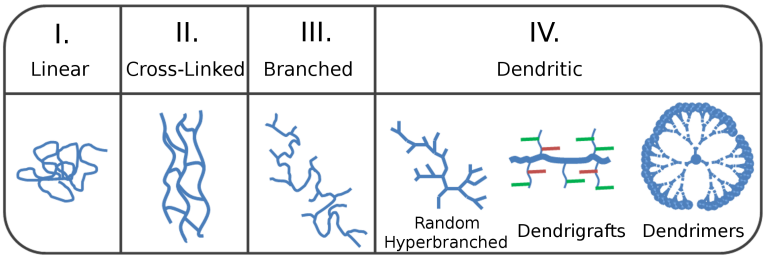


Figure 1.1: Schematic representation of the four major classes of macromolecular architecture. Adapted from [2].

An exact definition for *dendrimer* has been the subject of some debate and as such the term has evolved to a broader definition. It is now generally accepted that dendrimers are monodisperse macromolecular tree-like synthetic structures with a well-defined topology, composition, size and molecular weight [1, 6–10]. Indeed, this is what sets them apart from the other subclasses of dendritic polymers. The structural (synthetic) control over the topology of dendrimers is in sharp contrast to the statistical nature of the other subclasses of dendritic polymers, namely random hyperbranched [11] and dendrigraft polymers [12], where due to the synthesis methodologies and the nature of the branching groups used, the length of the branching chains is somewhat arbitrary, and hence, architectural asymmetries exist [1, 3, 13].

Dendrimers feature a unique core-shell molecular structure consisting of three basic architectural components: a core, the inner shells (generations) consisting of repetitive branches, and the periphery which is composed by the terminal groups (also known as the surface or outer shell) [1, 14]. Figure 1.2 illustrates these three regions using an idealized dendrimer.

Starting from a multi-functional core unit (denoted as G0), the struc-

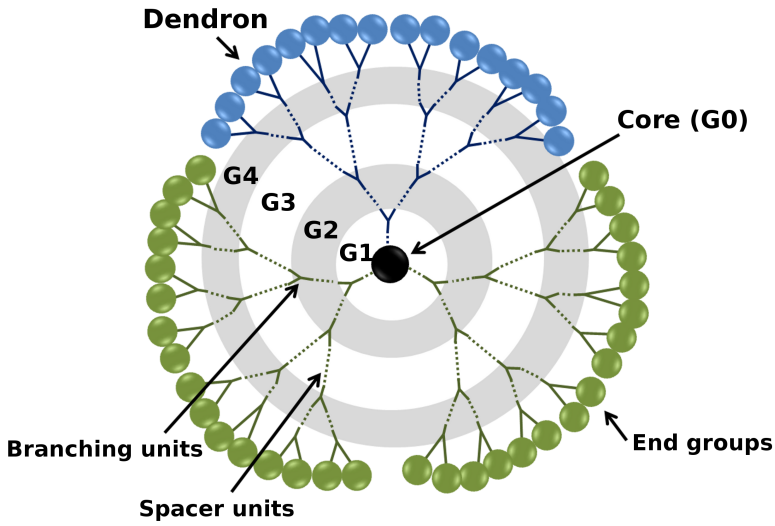


Figure 1.2: Schematic representation of a dendrimer composed by three dendrons. Adapted from [1].

ture branches in regular layers (shells) with each branching originating a new consecutive generation, composed by additional spacing and branching units. The multiplicity of both the core and the branching units determine the precise number of end groups, whereas mass amplification as a function of generation is also determined by the nature of the spacing units. Finally, the periphery may consist of either passive or reactive terminal groups. The branched structures linked in the form of segments to the core are known as dendrons [1–3, 7, 15].

The nature of these three architectural components determines the physicochemical properties found in different dendrimers, namely the size, shape, solubility, viscosity, reactivity, and conformational rigidity/flexibility [16–18]. Unique functional property patterns that depend on a dendrimer generation and are driven by its architecture are often found in these systems; such patterns are known as dendritic effects [19]. Within limits, appropriate *a priori* design of a dendrimer allows certain properties to be modulated or possibly tailored to requirement [6, 20, 21].

The widespread importance of dendrimers is due essentially to their unique architecture and the possibility of designing well-defined *func-*

tional macromolecules. A myriad of possibilities is available by incorporating different building blocks, whose functional groups can participate in chemical/physical processes [1, 22].

When referring to *functional dendrimers* it is common to distinguish between dendrimers incorporating specific chemically active functional groups (e.g. catalytically active, metal-coordinating or photoactive groups), and dendrimers whose functionalities serve mainly the purpose of controlling the characteristic chemical/physical molecular features (e.g. solubility, conformational flexibility) [16, 22, 23]. Since both fulfill a function, and has previously done by others [1, 23], the word "functional" is used here in its broader sense. It is important to emphasize that functional groups accounting for different functionalities can coexist in a single dendrimer, and that examples of placement of functional groups in any of the three architectural components aforementioned are available [6, 9, 23, 24].

Over the years, driven by the success of technological and medical applications employing dendrimers, the field has become a pervasive topic where elegant "tactics" developed by talented organic chemists have resulted in the construction of a broad range of dendrimers, from low generation ones to supramolecular assemblies [6, 24]. For example, recently Vötgle et al. have categorized dendrimers into twenty-nine distinct groups based on functional groups and chirality [1].

Having introduced the basic concepts, in the remainder of this section some historical background is presented. It merely intends to provide a quick overview on the evolution of the field, identifying key discoveries, and also, perhaps of greater importance to the reader, provide practical examples of different dendrimers, illustrating the versatility and synthetic potential of this architecture. An exhaustive enumeration of the types of dendrimers, their applications, and the available synthetic approaches is out of scope, but can be found in some excellent books [1, 2, 6, 14] and review articles [7, 10, 16, 17, 23, 25].

A particularly fascinating aspect of dendritic polymers lies in the fact that it all started as a puzzling conceptual problem, where theoretical predictions on the possibility of existence, and eventual characteris-

tics of such molecules, preceded their experimental "finding".

Indeed, in the early 1940s P. J. Flory (Nobel Prize in Chemistry, 1974) introduced the infinite (polymer) network theory, where he showed statistically that branched polymeric products are likely to appear in the polymerization processes [26–29]. His seminal work already included studies on tri- and tetrafunctional branching units [27, 28]. However, at the time, the experimental techniques (e.g. modern mass spectrometry tools) allowing the unequivocal characterization of such compounds did not exist. W. Stockmayer subsequently corroborated and extended Flory's work by developing a general formalism to estimate the most probable distributions of molecular sizes for certain types of branched-chain polymers [30]. It was only in the late 1970s that Vögtle et al. first reported the synthesis of several low molecular weight *cascade molecules* [31], using a repetitive growth synthesis relying on branching units. For the first time, *generational* molecules were prepared and characterized at each stage of the process; the result was a new core-shell macromoleculer architecture, now recognized as dendrimers.

Those initial studies were proceeded by several reports illustrating the usage of repetitive chemistry to prepare dendritic materials, including the noteworthy studies of Denkewalter et al. [32] describing the synthesis of polylysine–based dendrimers, i.e. the first peptide dendrimers (Figure 1.3). At the time, the asymmetric branching pattern and the presence of multiple chiral centers made those dendrimers unique.

In 1985 Tomalia reported the preparation of an entire series of poly-(amidoamine) dendrimers (PAMAM, generations 1 to 13) using a new synthetic methodology he named *divergent synthesis* (constructed from the inside out) [33]. The methodology, as well as the particular dendrimers synthesized, revolutionized the entire field. Tomalia's divergent approach allowed the synthesis of high generation dendrimers with good yields and purity, providing the first commercial route to dendrimers. Moreover, PAMAM dendrimers (Figure 1.4) are undoubtedly the most extensively characterized and best understood dendrimers, and the basis of most commercial applications involving dendrimers [5, 21].

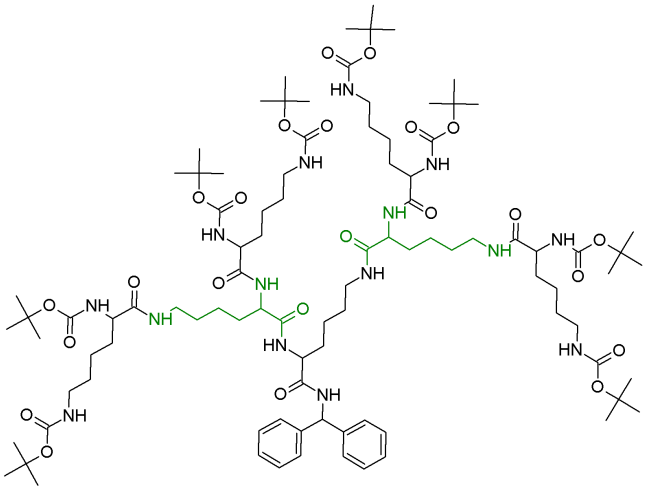


Figure 1.3: Second generation polylysine dendrimer [32].

It was also Tomalia who coined the term *dendrimer*[1]. It is derived from the Greek words *dendron* (tree) and *meros* (part of), underlying the tree-like nature of these molecules.

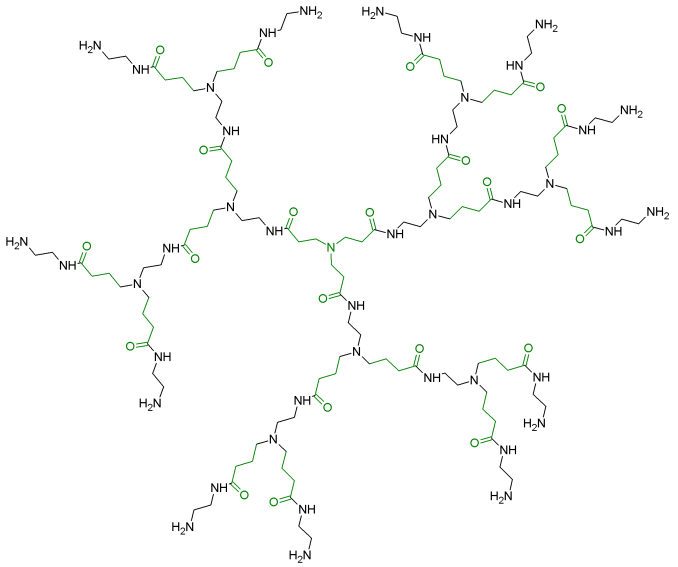


Figure 1.4: Second generation poly(amidoamine) dendrimer, PAMAM [33].

[1]Apparently inspired by his horticulturist hobby [2].

In the beginning of the 90s Fréchet and Hawker were responsible for a major development in the field: the introduction of the *convergent growth* approach for the preparation of dendrimers, i.e, built from the outside inwards [34]. They synthesized a series of poly(aryl ether) dendrimers which are also known as Fréchet-type dendrimers (Figure 1.5). Instead of expanding a core molecule *outwards* in a divergent fashion, the convergent growth starts at what will become the periphery of the molecule proceeding *inwards*, to afford multiple building blocks (dendrons) that are subsequently coupled to a core [2].

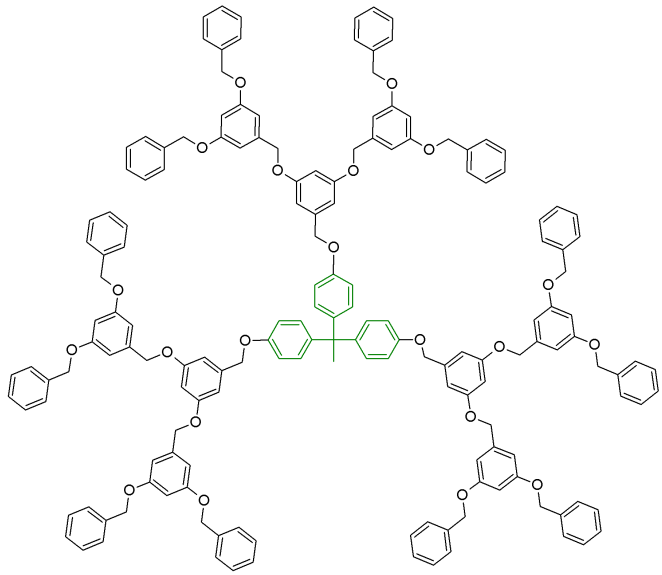


Figure 1.5: Third-generation poly(aryl ether) dendrimer, also known as Fréchet dendrimer [34].

1.2 Peptide Dendrimers: concepts, innovations and applications

As shown by Denkewalter's work with poly(lysine) dendrimers [32], from the early beginning of the dendrimer field there has been considerable interest in using amino acids as building blocks for dendrimer synthesis. The initial appeal was justified by the mere synthetic con-

venience of the existing peptide–coupling techniques [35] and the commercial availability of many proteinogenic and unnatural amino acids. Since those early studies, dendrimers containing amino acids have proven to display unique useful features justifying a wider interest.

Peptide dendrimers can generally be defined as dendrimers containing peptide bonds in their structure [10, 36, 37]. This broad definition includes dendrimers with an amino acid core, branching units, surface functional groups or any combination of the three, and is further broadened by taking into account all types of amino acids (naturally occurring and unnatural amino acids). Therefore, it is useful to classify peptide dendrimers into three different types based on the role played by the amino acids.

The first type includes dendrimers with a traditional (organic) dendrimer framework where amino acids compose only a specific architectural region (core, branches or surface). Examples include poly(hydroxyl)-surfaced PAMAM dendrimers where unnatural β-amino acids have been used as branching units [5], and poly(propylene imine) (PPI) dendrimers grafted with amino acids [38].

The second type encompasses dendrimers with amino acids composing the main framework (core and branches) while the periphery is grafted with other functional groups. Examples include peptide dendrimers functionalized with polyethylene glycol (PEG) [39] and dendrimers grafted with carbohydrate moieties (known as glycopeptide dendrimers) [6].

The third type corresponds to dendrimers composed exclusively by amino acids. Their tree-like structure is accomplished by using amino acids as core/spacer units and diamino acids as branching units (terminal amino acids act as surface functional groups). This group includes dendrimers which are essentially branching polyamino acids, such as poly(lysine) dendrimers, as well as more complex ones such as the well-known multiple antigenic peptides (MAPs) [36, 40], which contain a lysine scaffold bearing multiple copies of an antigenic peptide. In practice, the majority of peptide dendrimers referred in the literature and currently in use belongs to this type. Dendrimers of this type are also

the focus of the present thesis.

Amino acid-based dendrimers capitalize on the unique properties of the amino acid building blocks including chirality, hydrophilicity/hydrophobicity, biorecognition, and optical properties [6, 10, 37, 41]. Furthermore, the combination of a tree-like topology with the basic components of natural proteins results in molecules displaying protein-like properties and that are expected to adopt protein-like structure.

The possibility of taking advantage of peptide dendrimers intrinsic multivalency to graft other molecules, alongside with the ability of inducing specific microenvironments by combining different amino acids into their topology, makes peptide dendrimers extremely versatile molecules. For example, the distinctive internal composition created by the amino acid building blocks offers stereoselective sites for noncovalent host–guest interactions [10, 37, 42].

Also, in a sense, peptide dendrimers can be regarded as branched molecules incorporating peptidic segments. Peptides are ubiquitous in nature, playing an important role as antimicrobials, antivirals, analgesics, among many others [37]. Although in essence, peptides are built from amino acid monomers joined "simply" through peptide bonds (amide bonds), the diversity of functional groups that can be included through the side-chains of the different amino acids and the interplay between those groups leads to molecules exhibiting very different physico/chemical properties. The solubility and conformational flexibility of a peptide molecule is deeply correlated to the interactions that the side-chains of the residues can establish (e.g. hydrogen bonds, stacking interactions, or, when cysteine residues are present, disulfide bonds) and the nature of the functional groups present (e.g. hydrophobic or polar residues). The potential usefulness of a peptide and its efficiency also depend on its sequence. For example, peptides that have roughly the same number of amino acids can, depending on their sequence, penetrate or disrupt membranes [43].

It is the dual character of peptide dendrimers, dendritic and peptidic, that grants them their unique properties, which arise from the richness of the functional groups of peptides and the regular, ramified,

architecture of dendrimers. It is this duality, and the heterogenerational features it can encompass, that often allows low molecular weight peptide dendrimers (G2 or G3) to outperform the efficiency of other, more homogeneous, and higher generation dendrimers. This has recently been exemplified for DNA transfection [44] and bacterial biofilm inhibition [45].

Compared to their linear polymeric counterparts and to other (organic) dendrimers, the advantages of peptide dendrimers lie in the following features:

i. a protein-like structure that can act as a receptor by adapting to the shape of natural ligands [37];

ii. positive dendritic effects accounting for a cooperative amplification of function [46, 47];

iii. solubility in aqueous media without formation of aggregates [42];

iv. increased resistance to proteolysis, caused by their high degree of branching [37, 48];

v. biocompatibility, which can minimize their cytotoxicity [10, 37].

In 2005 Crespo et al. published an excellent review article surveying advances in the peptide dendrimer field; they finished the article by stating a crucial long-term goal:

> *The production of vast combinatorial libraries of dendrimers, based on the 20 genetically coded amino acids as well as hundreds of non-coded amino acids, with variations in branching units or building blocks at each generation can be imagined. Such combinatorial libraries should promote the discovery of lead structures and products with novel properties. Moreover, analysis of the libraries should facilitate exploration into new areas of application (...)*
>
> Crespo et al. *Chem. Rev.*, 2005, 105, 1663-1681.

In the past decade the group of J.-L. Reymond (University of Bern, Switzerland) successfully addressed this goal and disclosed an exciting

new approach to peptide dendrimers as artificial proteins [6, 42]. This approach involved essentially two main innovations.

(1) First the optimization of solid-phase peptide synthesis (SPPS) protocols to synthesize dendrimers with variable types, and amounts, of amino acids as spacer and branching residues. A divergent synthesis process under the standard conditions of SPPS [35, 49], with Fmoc-protected amino acid building blocks (Fmoc-SPPS), provided peptides with a dendritic topology when diamino acids were introduced in the sequences [6, 42]. A schematic description of the synthetic protocol is presented in Figure 1.6.

The side chains of the amino acids distributed throughout the branches can be used to provided specific functional groups (e.g. for catalysis) or be chosen for fine-tuning the physico/chemical properties of the dendrimers [41, 42].

With this protocol dendrimers with one to three amino acids between branches can be prepared with up to 37 amino acids (G1 to G3, MW~5×10^3 Da), Figure 1.7. This divergent synthesis was later combined with a convergent assembly approach that takes advantage of chloroacetyl to cysteine (ClAc) ligations to make larger peptide dendrimers (G4 to G6, MW~30×10^3 Da) [50].

(2) And second, the adaptation of the principles of "split-and-mix" combinatorial synthesis [51] to construct peptide dendrimers combinatorial libraries allowing the simultaneous screening of thousands of amino acid sequences [52]. For example, the screening of combinatorial libraries and optimization by amino acid sequence variations, resulted in the discovery of dendrimers with catalytic [53, 54] and ligand binding activities [55].

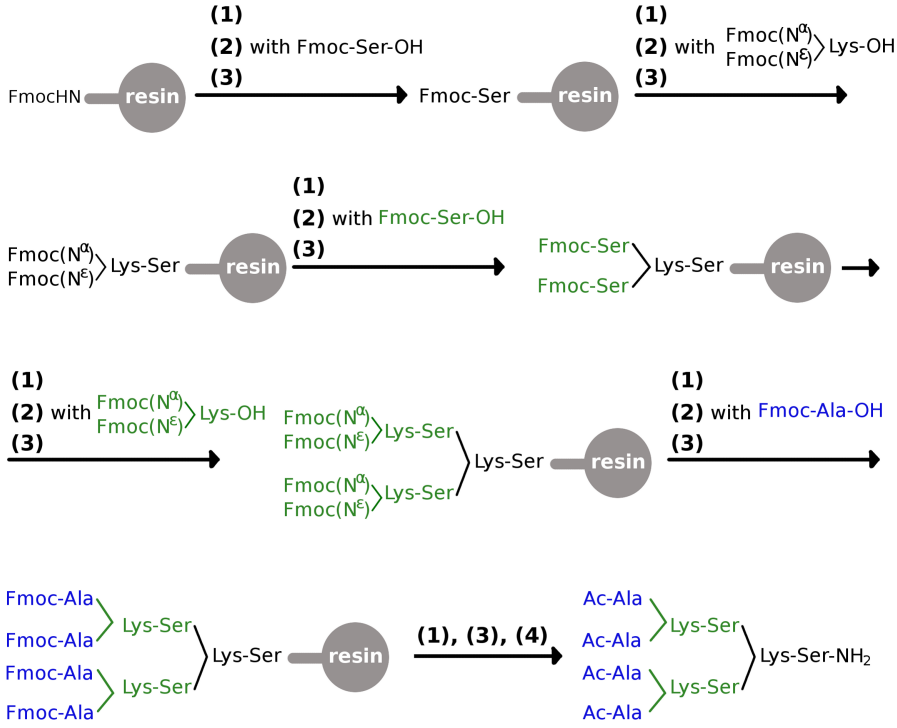


Figure 1.6: Solid Phase Peptide Synthesis (SPPS) of Peptide Dendrimers. The SPPS of peptide dendrimers using Fmoc-protected amino acids is exemplified using a dendrimer with sequence: $(AcAla)_4(Lys Ser)_2 Lys Ser\text{-}NH_2$. The dendrimer core is composed by C-terminal amino acids which are coupled first to the resin. Branching is promoted by the addition of diamino acids. Peptidic chains are extended by coupling N-protected amino acids (Fmoc-AA-OH) to the amino groups of the residues already attached to the resin. Each coupling step is followed by deprotection of the amino group to regenerate free N-terminus for the next coupling cycle. The final step of the synthesis involves the treatment of the solid support resin with acid (TFA) to remove side-chain protecting groups and cleave the peptide from the support. The cleaved material is then precipitated and purified by preparative reverse-phase high performance liquid chromatography (HPLC).

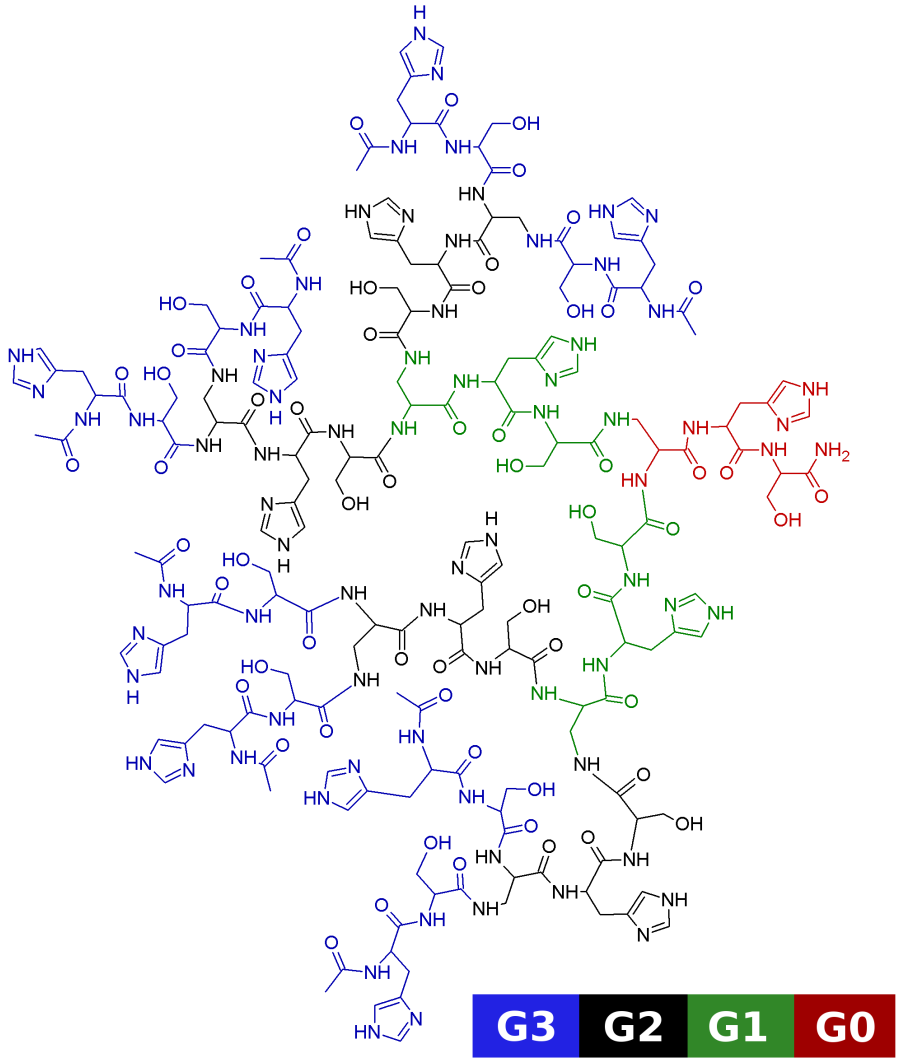


Figure 1.7: Third-generation peptide dendrimer with sequence: $(AcHisSer)_8(DapHisSer)_4(DapHisSer)_2DapHisSer\text{-}NH_2$ [46, 47]. The different generations are highlighted using a color code. *Dap* = *L*-2,3-diaminopropionic acid.

Table 1.1: Known applications of peptide and glycopeptide dendrimers.

Application	**References**
Ester hydrolysis catalysts	[46, 47, 50, 53, 57–60]
Aldol reactions catalysts	[50, 54, 61]
Labeling and drug delivery to cancer cells	[62, 63]
Models of metalloproteins: cobalamin ligands	[55, 64]
Models of metalloproteins: FE(II) ligands	[65, 66]
Pseudomonas aeruginosa biofilm inhibitors	[67–71]
Membrane disrupting antimicrobial dendrimers	[72–74]
Cell penetrating peptides	[44, 75, 76]

To date, the Reymond group has synthesized over 400 different peptide and glycopeptide dendrimers [42, 56]. Table 1.1 provides a summary of the most relevant dendrimers discovered and their applications.

Both the synthesis and applications of peptide dendrimers have recently been thoroughly reviewed elsewhere [6, 37, 41, 42, 52].

1.3 Structural studies of functional dendrimers

For experimentalists the topic of dendrimer molecular-level structure is a somewhat frustrating issue. On one hand there is wide curiosity in understanding the solution behavior of these molecules, and their interactions with other relevant molecules, on the other hand such desire stumbles on technical difficulties that most likely result from the intrinsic *functional* properties of dendrimers.

In many cases, standard nuclear magnetic resonance (NMR) techniques have been found to be of little use in determining structural defects, especially at higher generations, due to the "self-similarity" inherent to the dendritic architecture [1, 14, 77]. Also, approaches based on X-ray crystallography have not had any success; a fact that is normally rationalized as a consequence of dendrimers possessing a high structural flexibility with rapidly interchangeable configurations [77, 78]. In the case of X-ray structures, the sole exceptions found in the literature are the work of Bauer et al. [79] reporting the single-crystal structures of

a series of first generation polyphenylene dendrimers (which are essentially dendritic repetitions of aromatic rings), and the work of Michaud et al. [80] published in 2016, describing the structure of a second generation glycopeptide dendrimer co-crystallized with a lectin from *Pseudomonas aeruginosa*. The simplicity of those dendrimers, and the conformational rigidity observed, attest the difficulties found in this kind of studies.

Over recent years some success has been achieved with small-angle neutron scattering (SANS) [78], small-angle X-ray scattering (SAXS) [81] and PSGE diffusion NMR techniques [50, 55, 64, 74]. Nonetheless, those techniques provide only very rough pictures of the systems without any atomic-level detail.

Hence, it comes as little surprise that computational techniques such as molecular mechanics/molecular dynamics (MM/MD) have become a pervasive tool to explore the behavior of dendrimers, providing a sound basis for the molecular reasoning of the observed phenomena. Given the importance of such computational techniques to the present thesis, a more thorough discussion is presented in the next chapter.

To what pertains the application of computational methods to study dendrimers in general, suffices to say that almost every method as been (adequately) applied by different groups and even with different, and sometimes conflicting, findings. In analogy with the focus of experimental research, also the majority of computational studies have (and still do) focus mainly on PAMAM dendrimers.

The conformational features of PAMAM dendrimers in solution have been extensively studied using atomistic and coarse-grained models with different computational methods; including Brownian dynamics (BD) [82], Molecular Dynamics simulations [83], Monte Carlo simulations [84], and enhanced sampling techniques such as metadynamics simulations [85].

Since the first syntheses of dendrimers and until the mid-2005, almost all the computational work on dendrimers aimed at explaining the solution distribution of the different monomers. Such interest was triggered by a controversy resulting from the seminal work of de Gennes

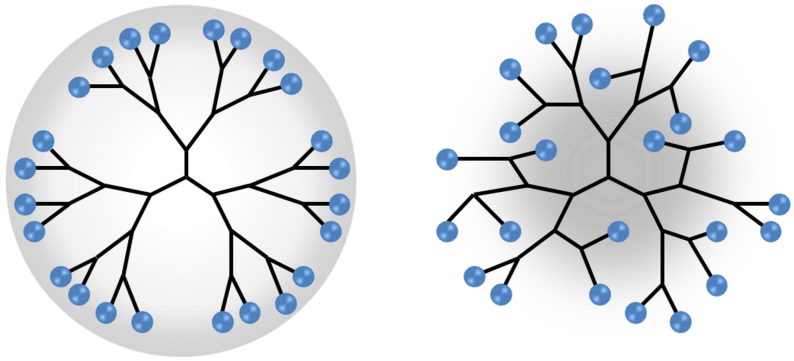


Figure 1.8: Segment density distribution of a hypothetical dendrimer according to the "dense-shell" (left) and "dense-core" (right) models. Adapted from [81].

and Hervet [86], which suggested that segments in successive generations are confined in spherical shells located at increasing distances from the dendrimer center of mass. This concept is known as the "dense-shell model" (Figure 1.8).

However, the currently accepted view on this issue is the "dense-core model". It was first proposed by Lescanec and Muthukumar [87] and latter confirmed by several SANS and SAXS studies with PAMAM dendrimers [1, 78, 81]. In the "dense-core model" maximum density is postulated to occur at the core of the dendrimer, with a decrease in segment density proportional to the distance from the core. The decrease in segment density towards the periphery is caused by partial back-folding of end groups into the interior of the dendrimer (Figure 1.8) [1].

Over recent years the focus of computational studies has shifted from monomer distribution models to solution effects (such as pH [88] and counter-ions [89]) and interactions with other molecules (e.g. host–guest [90], lipid membranes [91] and DNA [92]). Some excellent reviews surveying all the computational methods and findings resulting from their application to PAMAM and PPI dendrimers are available [77, 78, 90, 93].

Although there is ample computational work performed with PA-

MAM dendrimers, the same can not be said for peptide dendrimers. Indeed, the conformational preferences of such systems have remained mostly unaddressed; despite the multiple applications for such systems, the conformational variability that is to be expected due to the amino acid side-chains, and the humongous combinations of amino acid sequences accessible.

At the time we started the work presented in this thesis, only a few reports on the topic were available and all off them using MD simulations.

Dendrimers of the first type (organic framework grafted with amino acids, see the previous section) were studied by Cavallo et al. [94], who performed simulations of PPI dendrimers surface modified with phenylalanine. Their results suggested the self-inclusion of the peripheral amino acids into the dendritic inner shell.

Dendrimers of the second type, namely generations one to six of dendrimers with a poly(lysine) scaffold functionalized with other molecules were studied by Roberts et al. [95, 96], who observed a transition from a small molecule (flexible) behavior to a more polymer-like (dense-core) behavior with increasing generation.

Regarding peptide dendrimers composed solely by amino acids (third type) only three articles were available: two of those articles were published in 2009, namely one by Moiani et al. [97] addressing the specificities of the interaction between a second-generation dendrimer and a human immunoglobin; and a second one by Reymond and coworkers [98] showing that third-generation peptide dendrimers acquire a molten-globule like structure in solution. In both cases the simulation times used were short (10 ns and 20×4 ns respectively) and doubts regarding the equilibration of those systems remain. For the interested reader, a brief discussion specifically on equilibration issues in computational studies of dendrimers is available in reference [14], as well as a more general review [99] of such issues.

The third article was published by us in 2011 [100]. In that work, we performed multiple and long MD simulations to extensively sample the conformational preferences of five third-generation peptide den-

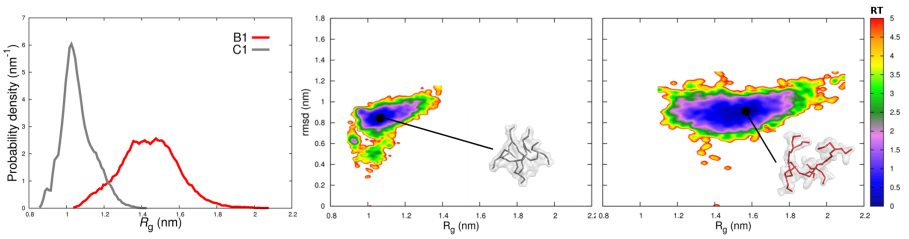


Figure 1.9: Radius of gyration histograms (left) and free energy profiles (middle and right) of two third-generation peptide dendrimers typifying the different conformational behavior observed by Filipe et al [100]. Dendrimer C1 (gray, middle): $(AcSerGly)_8(DapGluTyr)_4(DapHisThr)_2DapArgAla\text{-}NH_2$. Dendrimer B1 (red, right): $(AcGluSer)_8(DapGluAla)_4(LysAmbTyr)_2DapCysAsp\text{-}NH_2$. *Dap* = *L*-2,3-diaminopropionic acid. Amb = 4-aminomethyl(benzoic) acid.

drimers, including some known to bind aquocobalamine. Our study, specifically addressed the differences in the conformational preferences of peptide dendrimers with variable amino acid sequences.

The results underlined a high conformational flexibility, and two markedly distinct behaviors were found in dendrimers with the same topology but with different amino acid sequences, Figure 1.9. By resorting to energy landscapes and other analysis we showed that one of the dendrimers displayed mostly compact conformations clustered into distinct basins (C1, rough landscape), while the remaining dendrimers displayed mainly noncompact conformations with no significant clustering (B1, downhill landscape). These different structural preferences helped to explain why B1 is a stronger and faster aquocobalamin binder than C1.

Understanding the functional differences between peptide dendrimers that account for such entirely different conformational behaviors, is one of the main objectives of the present work. Over the next chapters of this thesis we report our endeavors and findings.

Meanwhile, two more computational studies done by other groups on peptide dendrimers have become available, namely: a thorough analysis of poly(lysine) dendrimers structural properties up to the fifth generation [101], and a study probing the interactions of polycationic an-

timicrobial peptide dendrimers with bacterial membranes [73].

1.4 Scope and Structure of this Work

Peptide dendrimers are the subject of this work. In the previous sections the basic concepts regarding these dendrimers were introduced. The reader can now have an overall idea of the state-of-art of the field and the challenges it faces.

As pointed out in the previous sections, peptide dendrimers applications have been extensively studied using experimental means. However, the underlying structural features of these systems have been, for practical reasons, somewhat neglected. This knowledge-gap limits the maturity of the field, and the broad acceptance of these systems.

Let us consider a simple first-generation peptide dendrimer with two spacer residues in each branch, i.e. a sequence of the form $(AB)_2bCD$, where b is a branching residue and A, B, C and D are amino acid residues which may or may not be identical.

If we restrict A, B, C, and D to be one of the 20 proteinogenic (natural) amino acids, and we also consider that b can be a single diamino acid, it becomes apparent that, even for such a simple system, there are 20^4 possible combinations of amino acids that could be used to construct such a peptide dendrimer. And we have not even considered the fact that B can actually correspond to two different amino acids, if an orthogonal deprotection scheme is used in the synthesis (e.g. Fmoc/Alloc chemistry). One can already start to imagine the myriad of distinct physical/chemical and structural features that are likely to result from the combination of different functional groups (e.g. side-chains of amino acids) in higher generation dendrimers. And, if one looks at the literature (which is summarized in the previous sections) it is noticeable that the most interesting and "applicable" peptide dendrimers are not this simple and are at least second-generations ones.

This exercise merely intends to show the enormous variability of possible dendrimer sequences and the need to define some general guidelines for selecting promising amino acid sequences. Even with combina-

torial libraries of amino acids, such as the ones described before, not all sequences can be attempted, and even such an approach would greatly benefit from knowing in advance some preferential dendrimer characteristics resulting from amino acids that are to be placed in specific positions.

Understanding in detail these systems, their interactions with other molecules, their behavior under different environmental conditions, the relation between the functional groups present in the amino acid sequence and the observed conformational features, are key milestones in the path leading to the development of truly knowledge-based peptide dendrimers; with all the benefits from such an approach, namely in terms of prospective applications.

Obtaining a solid fundamental knowledge of the key structural aspects of peptide dendrimers is the central goal of the work presented here.

We have taken advantage of the already available experimental information and have used molecular modeling methods to obtain the lacking atomic-level glimpse into these fascinating molecules world. In practice, we have performed a comprehensive set of computational studies using several dendritic systems and different methodologies. The results retrieved from our studies allow a new and deeper comprehension of these systems.

The basic concepts behind the computational methods employed are reviewed in the next chapter.

In the third chapter we describe our investigations on the conformational determinants of peptide dendrimer structure acquisition. Molecular dynamics simulations were used to study the influence of using different branching diamino acids, and also the effect of the placement of variable amounts of negatively charged residues in dendrimers with the same topology. This work entailed the development of several computational tools required to perform simulations of dendrimers. This work also originated a publication in *Macromolecules* [102].

The fourth chapter extends our previous studies to dendrimers with different amounts of positively charged amino acid resides. Again, molec-

ular dynamics simulations were used to investigate these systems, but the actual synthesis and experimental characterization of the dendrimers was also performed. An upgrade of the computational tools previously created was also needed to handle different (all-atom) force fields. This work confirmed a true structural dependence of peptide dendrimers on electrostatic effects, and has originated a manuscript that as been submitted for publication.

The pH-dependent conformational behavior of peptide dendrimers, and the interactions of catalytic dendrimers with substrate molecules are described in the fifth chapter of this thesis. Those topics were investigated using constant-pH molecular dynamics simulations. It should be emphasized that this is the first report on the usage of such methods to study any dendritic system, allowing us to reach a level of description of the systems that is absent in any of the other computational studies addressing pH-effects in dendrimers. This work involved adapting the in-house constant-pH MD software to allow for the simulation of dendritic systems. Furthermore, this work originated a manuscript that has recently been submitted for publication.

Finally, the sixth chapter summarizes the main aspects investigated, and discusses their repercussions and implications.

Chapter 2

Theory and Methods

Contents

Atoms are quantum entities and hence their behavior and interactions can in principle be adequately detailed in the framework of the quantum mechanics theory introduced at the beginning of the 20th century. However, due to the intrinsic complexity of atomic systems, alongside with several conceptual and practical problems, such an approach is virtually impossible for the great majority of relevant systems. Such difficulties are even more obvious when one considers biomolecular systems.

The complexity of molecular systems, which are essentially many-particle systems with multiple degrees of freedom, excludes the usage of analytical theoretical methods to characterize the static and dynamic features of such systems. As a consequence, numerical/computational simulations have been developed [103].

In a simplified view, one can consider that an adequate description of a molecular system through modeling and simulation involves essentially two aspects: first the construction of a physical model of the real system, which commonly includes several assumptions and approximations (i.e. a molecular model); second, a statistically meaningful way to determine the configurations accessible to the system (i.e. a sampling method).

Computer simulations of molecular systems have become of interest in the fields of chemistry, biology and biophysics, since they provide: (1) the means to rationalize and interpret experimental results from a microscopic point of view, (2) quantitative estimates of several properties of the systems and (3) the ability to study features of the systems under conditions that are beyond the reach of experimental techniques [104, 105].

In this chapter we shall describe the basic concepts of the computational methods relevant to the understanding of the following chapters. It is not an over detailed description and many important topics were left out. The detailed fundamentals of computer simulations of molecular systems and its applications, can be found in some excellent textbooks [103, 106–108].

2.1 Molecular Mechanics

The most common molecular models used in computer simulations are inspired either by quantum or classical mechanics. The distinction between different models lies in the level of detail used to represent the system, i.e. the degrees of freedom taken into account.

Molecular Mechanics (MM) theoretical models, also known as *force field methods*, follow from classical mechanics and consist in the usage of elegant "classical" functions to calculate the potential energy of a molecular system [104, 105, 109].

Such models implicitly consider two main assumptions. First, that the Born–Oppenheimer approximation is invariably assumed to operate, separating the electronic and nuclear motions. The electrons are regarded as presenting an endlessly smaller mass than the nuclei, and consequently they will rapidly adjust to any change in nuclear position. The second assumption is that the behavior of molecular systems, which are intrinsically of a quantum nature and whose motion is only fully described by the time–dependent Schrödinger's equation, can be accurately captured by the laws of classical mechanics.

In summary, in MM models, a molecular system is represented by a set of point masses where movement depends solely on the instantaneous coordinates of those point masses. Classical mechanics concepts, such as potential energies, partial charges or forces apply.

2.1.1 Force Fields

The total energy of a system can be represented by its Hamiltonian, $\mathcal{H}$. If $\mathcal{H}$ takes a classical mechanical form, than it can be decomposed in terms of potential energy contributions, U (that depend on the particles coordinates), and kinetic energy contributions, K (dependent on the particles momenta), as

$$\mathcal{H} = U + K \,. \tag{2.1}$$

For a closed system, the kinetic energy of a N-particle system, is given by [107, 108],

$$K = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m_i} \ , \tag{2.2}$$

where the masses and momentum of each particle are m_i and $\mathbf{p}_i$ ($i = 1, 2, ..., N$), respectively. As can be seen, dealing with the kinetic energy of the system is relatively straightforward. Defining the position–dependent potential energy is somewhat more complex.

In MM models, the mathematical form of U, the potential energy function, generally consists of several potential energy terms describing molecular connectivities and interactions. A textbook example of a potential energy function for a system of N atoms with position vectors $\mathbf{r}_i$ ($i = 1, 2, ..., N$) is [104]:

$$\begin{aligned} U(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = & \sum_{bonds} \frac{1}{2} K_b [b - b_0]^2 \\ & + \sum_{angles} \frac{1}{2} K_\theta [\theta - \theta_0]^2 \\ & + \sum_{dihedrals} K_\varphi [1 + \cos(m\varphi - \delta)] \\ & + \sum_{impropers} \frac{1}{2} K_\xi [\xi - \xi_0]^2 \\ & + \sum_{pairs(i,j)} \left[\frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^{6}} \right] \\ & + \sum_{pairs(i,j)} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{\varepsilon_r r_{ij}} \ . \end{aligned} \tag{2.3}$$

The potential can be separated into bonded and non-bonded terms (see below). In this functional form, covalent interactions are taken as summations over 1-2, 1-3 and 1-4 bonded terms, while non-bonded interactions are pairwise additive, modeled by sums over Lennard-Jones and Coulombic potentials.

The designation *force field* (FF), usually refers to both the potential energy function and the corresponding parameters used (K_b, b_0, K_θ, etc.). In other words, force fields convey molecular interactions in a

quantitative way, employing a set of equations, estimates of constant parameter values and free parameters. Consequently, force fields are empirical [105, 109, 110].

To a large extent, even the form of the potential energy function is empirical, since it depends on considerations of what terms are and are not required to effectively model the underlying quantum mechanical phenomena using a simple multi-term classical depiction. The form of Equation 2.3 represents a compromise between simplicity and accuracy. Given the empirical nature of potential energy functions, FFs with additional terms, or that account for coupling between different types of interactions are common. For example, some FFs contain mixed terms as $K_{b0}[b - b_0][\theta - \theta_0]$, which directly couple bond-length and bond-angle vibrations [104, 110]. The actual potential energy functions used in Chapters 3 to 5 of this thesis are presented in Appendix A (page 127).

Force fields are also very distinct in the way empirical parameters are determined. In some force fields, such as AMBER [111, 112] or OPLS [113], the parameters are fitted to high-level *ab-initio* quantum calculations of small molecules. A different philosophy is followed in force fields such as GROMOS [114–116], where parameters are obtained from fittings to experimental data (e.g. crystal structures, X-ray data of small molecules, free energy of solvation, among others).

Another aspect that further differentiates the existing force fields is the representations of the particles that compose the system. Some explicitly depicted all the atoms in the system (known as *all-atom* force fields), however, coarse–graining of some atoms or even of entire monomers (e.g. amino acids) is common practice since it allows for significant computational savings. For example, in force fields of the GROMOS family, non-polar hydrogens (aliphatic and aromatic) are typically subsumed into the atoms they are covalently bonded; the parameters of those atoms are modified accordingly. These force fields are called *united–atom*.

Force fields define the internal energy of a system, and ultimately the quality of a simulation will result from the ability of the force field chosen to accurately described the phenomena one is interested in.

2.1.2 Bonded Interactions

The first four terms of Equation 2.3 describe bonded interactions. Bonded terms represent the displacements of bond lengths (stretching), angles (bending) and rotations about covalent bonds (torsional angles) [104, 110].

The first term in the equation, *bonds*, is a harmonic potential accounting for the covalent bond stretching along bond b. The associated force constant (K_b) and minimum–energy bond length (b_0) depend on the particular type of bond.

The second term in the equation, *angles*, is also a harmonic potential but for three-body interactions and represents the bending of bond angles. A force constant (K_θ) and a bending–angle reference value (θ_0) must also be defined.

The third and fourth terms in Equation 2.3 both describe torsional angle interactions (four-body). The third term, *dihedrals*, is a sinusoidal potential for flexible dihedral angles φ, which can make 360° turns and have multiple dihedral minima. The periodicity of the torsions is described by the parameter m (multiplicity) whereas δ stands for the phase offset.

The fourth term, *impropers*, is a harmonic term applied to ξ dihedral angles in order to disallow transitions. Such terms are normally applied to dihedral angles of atoms in aromatic rings in order to maintain the planarity of the ring, or to chiral groups in order to maintain the chirality [104, 109].

2.1.3 Non-bonded Interactions

The last two terms of Equation 2.3 describe non-bonded interactions. Both terms are sums over all pairs of atoms in the system, and therefore, these interactions can be intra- or inter-molecular.

The fifth term of the equation represent the potential energy associated with van der Waals (vdW) interactions. It is common to model such interactions with a Lennard-Jones 12-6 potential (as in Equation 2.3) which contains a $1/r^6$ attractive term and a $1/r^{12}$ repulsive term.

At short distances the vdW interaction energies are very high, accounting for the repulsion between the particles. As the distance between the particles increases it is the attractive part of the function that becomes preponderant, but the magnitude of the interaction decreases. Therefore, vdW interactions are considered as short-ranged and it is common to reduce the calculation of the interaction term to pairs of atoms that are within a *cutoff* distance [103, 106, 109].

vdW interactions are a generic way of accounting for specific non-bonded forces, which include the forces between permanent dipoles, permanent and corresponding induced dipoles and also London dispersion forces (between two instantaneously induced dipoles).

Electrostatic interactions occur between groups that bear formal charges, or that contain substantial partial charges. In the last term of Equation 2.3, Coulomb's law is used to account for electrostatic interactions. Such interactions occur between non-bonded atoms i and j with charges q_i and q_j at a distance r_{ij}. ε_0 is the relative permittivity of vacuum and ε_r the relative permittivity of the medium[1].

Electrostatic interactions are particularly relevant in simulations of biomolecular systems since the interaction energy associated to them does not show a rapid decay with r_{ij} (proportional to $1/r_{ij}$) as in the case of vdW interaction. Therefore, electrostatic interactions are considered to be long-ranged and there is a conceptual need to account for their effect beyond the system boundaries.

Computationally the calculation of these long-ranged electrostatic interaction is extremely cumbersome, and stratagems allowing for significant computational savings have been developed. Typically electrostatic interactions are evaluated using either continuum or Lattice-sum methods.

In continuum methods, as for instance the reaction-field approach, the real charge distribution beyond a cutoff distance is treated using a mean-field approximation represented by a dielectric constant ε_{RF},

[1]Depending on the force field and the conditions it aims at describing (gas- or condensed-phase), this value can be a predetermined constant or a distance-dependent function. For example, GROMOS force fields use $\varepsilon_r = 1$, since they were developed to simulate biomolecules in explicit aqueous environment.

and an average ionic distribution I_{RF}. Electrostatic interactions between atom pairs within the cutoff are treated explicitly [104, 109, 117].

In Lattice-sum methods, such as the Ewald summation, and all its fast Fourier transform variants (e.g. P^3M and PME), the system is represented as a neutral central unit cell that is infinitely repeated in all dimensions,[2] and the electrostatic interactions between all pairs of atoms (including periodic images) are summed. The neutrality condition and the breaking of the calculations into real and reciprocal space makes the energy summation convergent [118, 119].

2.2 Sampling Methods

Under physiological conditions molecules are not isolated, they interact with other molecules such as solvent ones. Such interactions lead to conformational fluctuations (e.g., entropic effects that result from collisions at non-zero temperatures) and hence, the thermodynamic behavior of molecules can only be described if we take into account this dynamics. This means that a system is not characterized by the energy of a single conformation, but by its free energy [120, 121].

The entire behavior of a dynamical system can be conceptualized in terms of a multidimensional space defined by all the values of position ($\mathbf{r}$) and momenta ($\mathbf{p}$) that are accessible to it.[3] For a system composed of N atoms, this space, which is known as *phase space* (represented as $\Gamma(\mathbf{r}^N, \mathbf{p}^N)$) will be defined by $6N$ values; three coordinates per atom and three components of the momentum. Under equilibrium conditions, the regions of this space will be populated with different probabilities in accordance with a Boltzmann distribution (i.e. the probability density function governing the populations of the phase space) [106, 108, 122].

The systems that are currently tractable with computational simu-

[2] This treatment of the system boundaries is known as periodic boundary conditions (PBC) and is not an exclusive feature of Lattice-sum methods, but rather a general way to deal with the boundaries of systems that are periodic or contain explicit solvent molecules. PBC can also be used in simulations with the continuum methods mentioned in the text.

[3] In biomolecuar systems not all arbitrary combinations of $\mathbf{r}$ and $\mathbf{q}$ are accessible.

lation methods have a very large number of atoms (10^4-10^6 or more)[4], implying that the underlying phase space can be huge. It is thus, normally impossible to describe the entire phase space of a system. Alternatively, simulations of molecular systems focus on the generation of a statistically representative set of configurations at well defined bulk thermodynamic properties (e.g temperature, pressure, etc.), a so-called *ensemble*.

Herein, we shall describe two methods that provide a correct Boltzmann–weighted sampling of ensembles: Molecular Dynamics (MD) and Monte Carlo (MC) simulations.

2.2.1 Molecular Dynamics (MD) simulations

One way to generate an ensemble of configurations it to try to mimic the true temporal evolution of the system by solving Newton's equations of motion over time. This is known as the Molecular Dynamics (MD) method [105, 106].

Newton's law of motion (or second law) elaborates the relation between a particle's mass (m), acceleration ($\mathbf{a}$) and an applied force as: $\mathbf{F} = m\mathbf{a}$. Since acceleration is the second derivative of a position vector - $\mathbf{r}$ - with respect to time (and the first derivative of velocity)[5] we can rewrite Newton's law for a system of N-particles as:

$$\frac{\mathbf{F}_i}{m_i} = \frac{\partial^2 \mathbf{r}_i}{\partial t^2} \tag{2.4}$$

with $i = 1, 2, ..., N$. The force, $\mathbf{F}_i$, acting on each particle of the system is determined by the gradient of the potential energy, U (see Equation 2.3), relative to the position of that particle, as:

$$\mathbf{F}_i = -\frac{\partial U(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)}{\partial \mathbf{r}_i} \tag{2.5}$$

where U is a scalar quantity, that depends on the positions of all the

[4] A number that is still very small when compared with Avogadro's number, that is, macroscopic sizes.

[5] We could equivalently write this with respect to velocity as: $\frac{\mathbf{F}_i}{m_i} = \frac{\partial \mathbf{v}_i}{\partial t}$.

particles in the system and the force field [105, 107]. Once the forces (vectorial quantity) acting on all the particles are calculated through Eq 2.5 we can use Eq. 2.4 to obtain the particles new positions and velocities. However, this can only be achieved using numerical methods (many-body problem) that discretise time and solve the equations via a finite difference method. Those algorithms are called MD *integrators* and textbook examples include the Verlet and Leap-frog schemes [104, 105].

Normally, the equations are integrated over short time steps, Δt (1 to 2 fs for all-atom simulations). At each step, the forces on the atoms are computed and combined with the current positions and velocities to generate a new set of positions and velocities a short time ahead. The force acting on each particle is assumed to be constant during the time step of each iteration [108]. This is easily understood by looking at the equations used in the Leap-frog algorithm [123]; the velocities of a particle ($\mathbf{v}_i$) at time $t + \Delta t/2$ are defined by:

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{\Delta t}{m_i}\mathbf{F}_i(t), \tag{2.6}$$

and the positions at instant $t + \Delta t$ by

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right). \tag{2.7}$$

Usually MD simulations are run for tens to hundreds of nanoseconds.[6] The required duration of a simulation can only be decided by evaluating the convergence of the properties one wishes to study (which is very system-dependent).

The final result of an MD simulation is a trajectory (configurations as a function of time) of the molecular system, which can be "viewed" as a sequence of points in phase space that are connected in time. Average as well as time-average information can thus, be obtained [105, 106, 108, 121].

[6] A 1 ns simulation using Δt =1 fs corresponds to 10^6 steps.

2.2.2 Monte Carlo (MC) simulations

When one is not interested in the temporal properties of a system, Monte Carlo (MC) simulations can be used to sample the different states of a system in a stochastic manner [106, 124]. In MC simulations, the configurations sampled are not time-correlated as in MD. Each new configuration depends only on the preceding configuration.

The usual approach is to use a MC algorithm together with a Metropolis acceptance/rejection criterion (Metropolis MC) [125]. Starting from a configuration of the system $\mathbf{r}_S$, a new configuration $\mathbf{r}_{S+1} = \mathbf{r}_S + \Delta\mathbf{r}$ is generated by random displacement of one (or more) particles ($\Delta\mathbf{r}$). Then, the change in potential energy between $\mathbf{r}_S$ and $\mathbf{r}_{S+1}$ is evaluated (ΔU).

According to the Metropolis criterion, the probability of acceptance, p is:

$$p_{S\rightarrow S+1} = \min\left[1, \exp\left(-\frac{\Delta U}{k_B T}\right)\right] \quad . \tag{2.8}$$

The new configuration is always accepted if it has lower potential energy than the previous one ($\Delta U \leq 0$). However, if $\mathbf{r}_{S+1}$ has higher energy than the previous state ($\Delta U > 0$), a random number, *rand*, between 0 and 1 is generated. If the value of $rand < p_{S\rightarrow S+1}$ the move is also accepted. This procedure has the effect of allowing the proper sampling of high energy states.

Upon acceptance, the configuration becomes part of the productive ensemble and is used as starting point for the subsequent random displacement ($\mathbf{r}_{S+2} = \mathbf{r}_{S+1} + \Delta\mathbf{r}$). If on the other hand the criteria are not met, $\mathbf{r}_{S+1}$ is rejected, $\mathbf{r}_S$ is counted again and a new random displacement starting again from the initial configuration ($\mathbf{r}_S$) is attempted.

There is no momentum component in usual Monte Carlo simulation and such simulations sample from a $3N$-dimensional phase space corresponding to the positions of the atoms. This also implies that kinetic features of the systems can not be studied using this method (see Equation 2.2).

The random displacements of particles do not need to be Cartesian moves, dihedral or internal coordinate moves, among many others are

also possible [124]. Moreover, usage of the overall formalism of the MC method described here is not restricted to sampling of molecular configurations. Indeed, it can be applied to sample the protonation states of biomolecules; in that case instead of random displacements of atoms we will have (de)protonation of titrable groups and, instead of ΔU it is the free energy associated with the protonation reaction that needs to be evaluated (more details in the next sections) [126–128].

2.3 The microscopic and macroscopic link: statistical mechanics

The ensemble resulting from proper sampling of phase space is characterized by a set of microstates (configurations) but also by a set of bulk macroscopic thermodynamic quantities (known as *thermodynamical boundary conditions*) to which the collection of microstates is directly related [106, 129, 130]. For example, in MD simulations, the equations of motion for a system are the ones being integrated and thus, a microcanonical (constant total energy) ensemble is sampled by default. One can think of the resulting microstates as being sampled from phase space along a contour of constant energy.

The thermodynamical boundary conditions used, together with their reference–macroscopic values, define the thermodynamical ensemble that is sampled during a simulation. Normally simulation methods employ one of four thermodynamical ensembles: microcanonical, canonical, (semi-)grand-canonical, or isothermal-isobaric. A list of thermodynamical ensembles is provided in Table 2.1.

The thermodynamical boundary conditions involving extensive quantities (N, V, E) should be satisfied exactly at any point of the simulation, whereas those involving intensive quantities (μ , P, T) should fluctuate around the defined average value.[7]

[7]For the sake of clarity, the *IUPAC Gold Book* definitions for extensive and intensive quantities is provided. *Extensive*: a physical quantity that is the sum of the properties of separate non-interacting subsystems that compose the entire system. *Intensive*: physical quantity whose magnitude is independent of the extent of the system.

Table 2.1: The thermodynamical ensembles. N and μ are respectively the number and chemical potential of all particles for all species; V the volume; P is pressure; E is energy; H, L and R are thermodynamic potentials related with energy (E), namely the enthalpy ($H = E + PV$), Hill energy ($L = E - \mu N$), and Ray enthalpy ($R = E + PV - \mu N$); T is the temperature. Note that the generalized ensemble is not a physical ensemble since its size is not specified (no extensive boundaries). Adapted from [130].

Themodynamical Boundaries	Ensemble
NVE	Microcanonical
NVT	Canonical
NPH	Isoenthalpic-isobaric
NPT	Isothermal-isobaric (Gibbs)
μVL	Grand-microcanonical
μVT	Grand-canonical
μPR	Grand-isothermal-isobaric
μPT	Generalized

Unfortunately, the microcanonical ensemble resulting from a standard MD simulation does not correspond to the conditions under which most experiments are performed; it is often desirable to sample configurations from a canonical or isothermal-isobaric ensemble instead. This can be achieved by introducing specific modifications in the system Hamiltonian or in the equations of motion to maintain a constant temperature or pressure [107, 108, 130]. A description of the thermostat and barostat algorithms used in computer simulations is out of scope, but can be found in an excellent review article by P. Hünenberger [130].

It must be noted that in usual MM/MC simulations it is the canonical (NVT) ensemble that is sampled by default, since the algorithm does not involve momenta or kinetic energy and a reference temperature is introduced through the Boltzmann factor in Equation 2.8. Modified versions of the MC procedure allowing for simulations in the microcanonical and grand-canonical ensembles are available [130].

The connection between the microscopic behavior and macroscopic properties of molecular systems is governed by the laws of *statistical mechanics* which allow us to express thermodynamic (macroscopic) prop-

erties in terms of microscopic quantities [105, 122, 131]. Using methods such as MD or MC there are a number of properties that can be immediately obtained, such as the velocities or positions of the atoms. However, such information cannot be compared with experimental data because none of the available experimental techniques directly provides such detail. In fact, a typical experiment determines an average quantity, averaged over all the particles in the system during the time needed to perform the measurement.

The basic assumption underlying statistical mechanics is the *ergodic hypothesis*, which states that in equilibrium, following a single molecule for a sufficiently large time interval, is identical (in terms of sampling of the system) to capturing the individual features of a very large ensemble of identical molecules at a single instant [99, 132]. This means that, in equilibrium, averages taken along "sufficient" time should be identical to ensemble-averages. Then, the average value of a certain thermodynamic macroscopic property, A, obtained for instance from a MC simulation, is simply the arithmetic average,

$$\langle A \rangle = \frac{1}{M} \sum_{M} A(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) \ , \tag{2.9}$$

where $\langle ... \rangle$ denotes an ensemble average and M is the number of configurations sampled in the simulations [108]. Also, as the ensemble accounts for every potential microstate the system can visit, s, and contains the relative number of times each microstate is visited (indistinguishable configurations), $n(s)$, we can express the average of A in function of the probability of the occurrence of each microstate, $P(s) = n(s)/M$, as,

$$\langle A \rangle = \sum P(s) A(s) \ , \tag{2.10}$$

where $A(s)$ is the value of A for state s. As a consequence, we observe that all that is needed to adequately describe the system is a way to determine the weighted probability of each microstate, and that is exactly what the MD and MC methods will provide in the limit of perfect sampling. For example, under conditions of constant number of particles, pressure and temperature (an isothermal-isobaric ensemble) the proba-

bility density is given by the Boltzmann distribution, as

$$P(s) = \exp(-(E_s + PV_s)/k_B T)/Q \tag{2.11}$$

where E_s and V_s are the energy and volume of the microstate. Q is the *partition function*, it is used to normalize the probabilities of each microstate such that the sum of the probabilities equals one. All thermodynamic quantities can be expressed in terms of the partition function. Some, like Gibbs free energy ($G = -k_B T \ln Q$) directly depend on it, whereas others, like internal energy or heat capacity are related to derivatives of the partition function [108, 122].

If configurational sampling is done correctly, with properly weighted microstates and sufficient sampling then we can compute accurate averages, fluctuations and other statistical values for many types of properties, including structural features, thermodynamic quantities and even time-related dynamical properties (if the simulations are performed with a time-dependent method). However, "normal" MD and MC simulations preferentially sample the lower-energy regions of phase space due to the magnitude of energy barriers and the time-scale of the configurational fluctuations. Consequently, entropic thermodynamic quantities that are directly related to the partition function (e.g. entropy, free energy), and therefore require an adequate sampling of high-energy regions, are normally poorly converged and inaccurate in any real simulation (i.e. in practice a true ergodic trajectory in never achieved) [99, 105, 106, 108].

Excellent books explaining the concepts of statistical mechanics have been published, for example, by D. McQuarrie and D. Chandler [122, 129].

2.4 Modeling titration states

The following sections will focus on computational techniques to study the electrostatic properties and pH effects on peptidic molecules. It is convenient to describe such techniques and the underlying concepts from the point of view of proteins, for the simple reason that the topics covered were essentially studied in the context of proteins. Nevertheless, all the topics discussed can be used in general with any peptidic system (like a peptide dendrimer).

Also, there are two types of titrable ionizable groups, protonatable and redox. We will focus on protonatable groups, but again, the formalism described can be used with small modifications to model the binding or release of electrons in redox groups [133, 134].

Titrable groups in peptidic systems (side chains of acidic and basic amino acids plus non-capped N- and C-termini) participate in acid-base reactions binding or releasing protons, i.e. protonation/deprotonation reactions. Therefore, titrable groups can adopt charge states that influence the stability, solubility, catalytic features, molecule–molecule interactions, basically every relevant aspect that concerns peptides in real situations [135–137]. Given their importance it is only natural that computational techniques have evolved in order to understand and try to predict the titration behavior of such groups.

If a single titrable group is present in solution, its protonation equilibrium can be characterized by its proton binding affinity using the formalism of acid-base equilibrium constants. The acid-base reaction

$$\mathrm{AH} \rightleftarrows \mathrm{A}^- + \mathrm{H}^+$$

is characterized by a equilibrium constant, K_a, that determines the strength for the dissociation of the acid into its conjugated base and a proton:

$$K_\mathrm{a} = \frac{[\mathrm{A}^-][\mathrm{H}^+]}{[\mathrm{AH}]} \ , \tag{2.12}$$

with $\mathrm{p}K_\mathrm{a} = -\log K_\mathrm{a}$. By tacking negative logarithms on both sides of Equation 2.12 and rearranging the terms, one obtains the familiar

Henderson–Hasselbalch equation

$$\mathrm{pH} = \mathrm{p}K_\mathrm{a} + \log \frac{[\mathrm{A}^-]}{[\mathrm{AH}]} = \mathrm{p}K_\mathrm{a} + \log \frac{f}{1-f} \ , \tag{2.13}$$

where f denotes the fraction of deprotonated groups. Since there are no other titrating groups, this conventional $\mathrm{p}K_\mathrm{a}$ expresses the protonation equilibrium at each pH value (in a sense it is pH-independent) and its value is equal to the pH when $f = 1/2$ [137]. It becomes obvious that, in order to have a single titrating group (e.g. amino acid), the tabulated $\mathrm{p}K_\mathrm{a}$ values found in any chemistry textbook refer to the group alone in solution or to a model compound of the group ($\mathrm{p}K_\mathrm{mod}$) [138, 139].[8]

Relevant peptidic systems normally have more than one (simultaneously) titrating group and the interactions between the groups make the "mid-point titration $\mathrm{p}K_\mathrm{a}$" of each group depend on the pH of the solution. Multiple experimental examples of such $\mathrm{p}K_\mathrm{a}$ shifts have been reported [135].

The charge state and titration profile of a particular group depends on the surrounding environment. Several interactions can contribute to alter the charge state of a group: electrostatic interactions with charges of other titrable groups or partial charges of non-titrable groups in the same molecule; location of the group in the molecule (internalized vs. solvent exposed); changes in the molecule conformation states; and the pH of the solution [137]. pH is of paramount importance since the direct result of a pH change is a modification in the equilibrium concentrations of the protonated and deprotonated forms of titrable groups [140].

The $\mathrm{p}K_\mathrm{a}$ value of a single titrable site in an otherwise neutral protein in known as the *intrinsic* $\mathrm{p}K_\mathrm{a}$, $\mathrm{p}K_\mathrm{int}$. Even in a protein, if a single titrable site exists, then the protonation equilibrium is given by Equation 2.13. However, and as explained before, when other titrable groups exist, the interactions between them (site–site coupling) need also to be considered. This means that the actual protonation profile of each titrable site

[8] $\mathrm{p}K_\mathrm{mod}$ values result from model compounds, which are chemical analogues of the titrable site in solution. The analogue should in principle mimic the chemical properties of the protein environment (e.g. capped pentapeptides with a single, central, titrable amino acid).

will result from cumulative site–site cooperative effects but also, that it will depend on pH.

It is still possible to express the protonation equilibrium of each site through an *effective* pK_a, pK_{eff}, which is defined by a mathematical expression identical to Eq. 2.13 were pK_a is replaced by pK_{eff}. However, pK_{eff} is a pH-dependent quantity and using a single pK value to characterize the protonation equilibrium of a site becomes inappropriate.[9]

In summary, the concepts behind traditional acid-base equilibrium alone are insufficient to study the titration of proteins.

2.4.1 Continuum Electrostatics

Understanding the properties of aqueous solutions requires models of the solute, the solvent, and the interactions between them [141]. As previously mentioned, electrostatic interactions differ from other molecular interactions by virtue of their non-local character. For instance, bond lengths are noticeably local, van der Waals interactions vanish after short atom–atom distances, but electrostatic interactions are long-ranged.

There are essentially two routes to model electrostatic interactions, the previously described MM models and *Continuum Electrostatic* (CE) approaches.

Continuum electrostatic (CE) models provide a (simplified) macroscopic–like level of description of a system. In contrast to the microscopic atomic-level view of MM models, in CE models, solute and solvent are described in terms of average values. The charge densities and dielectric properties change continuously in space, with the solute (protein) and solvent being regarded as different media with low and high dielectric constant, respectively. A schematic description of a CE model for a protein system is shown in Figure 2.1.

In CE models, charge density, ρ, is simply the average charge that arises at a particular point as a result of the microscopic charge distribution caused by atomic movements. A solute charge density consists

[9] Notice that even in this situation one can still define a mid-point pK_a, which is the pH value when $f = 1/2$; the so-called pK_{half}.

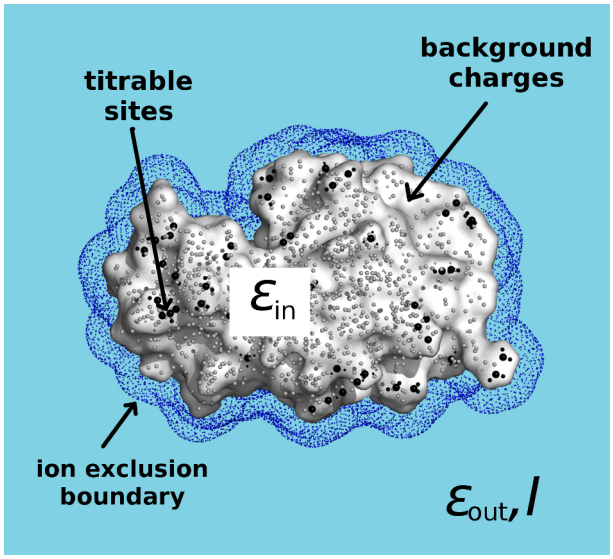


Figure 2.1: Continuum electrostatic model of a protein in solution.

of the partial charges of all the atoms and bounded ions (formal charges can also be used). A distinction between charges of titrable groups and other charges of the solute (background charges) is often used. When present, the distribution of counter-ions will be influenced by the solute charge states. To model the maximum proximity between counter-ions and the solute an ionic boundary is normally included. This ion-exclusion layer accounts for the finite size of ions. The overall counter-ion distribution is determined by the electrostatic potential and the solution ionic strength, I. The charge contribution of free solvent molecules is on average zero [140, 142].

The empirical dielectric constants (ε_{in} for the solute and ε_{out} for the solvent) try to account for all the contributions that are not explicitly defined in the model. Dielectric constants are phenomenological parameters that give an average measure of the systems polarizability by an electric field. Both electronic and orientational polarization effects are represented by the continuum dielectric constants. The higher the value of ε, the greater the capability of configurational rearrangements in response to an electric field [143–145]. The solvent dielectric constant is expected to capture the dipolar effects arising from an asymmetric

charge distribution in the solvent molecules (permanent dipoles); also, if the solvent is water, there is a high dipolar rotation and a high dielectric constant should be used (typically $\approx$ 80) [141, 144]. Furthermore, the shielding effect of the solvent is implicitly reflected in the high ε_{out} [141, 144, 146].

In the case of proteins (as solutes), their interior is normally modeled with a low dielectric to account for the orientational polarizability resulting from fluctuations of atomic charges around equilibrium positions [147]. The boundary between solute and solvent dielectric regions is usually defined by the solute solvent-accessible surface.

One of the main issues regarding the usage of CE methods is the choice of the solute conformation, since calculations using different conformers may yield different results. CE methods can adequately capture the electrostatic features of rigid structures but alternative methods should be used to sample relevant conformers [127, 139].

2.4.2 The Poisson-Boltzmann equation

Having defined the overall features of the CE model, the charge distribution, $\rho(\mathbf{r})$, and dielectric constant, $\varepsilon(\mathbf{r})$, can be assigned to each point, $\mathbf{r}$, in space and the dimensionless electrostatic potential, $\phi(\mathbf{r})$, can be determined by solving the Poisson equation for inhomogeneous media:

$$\nabla \cdot [\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \ . \tag{2.14}$$

If $\varepsilon(\mathbf{r})$ is uniform throughout space and there are no mobile ions in solution, then Equation 2.14 is reduced to the familiar Coulomb's law [141, 146].

The first use of CE methods goes back to continuum solvation models, namely to the Born model of ionic solvation. In that model, the (Gibbs) free energy of solvation is regarded as the electrostatic work accounting for the transfer of an ion (represented simply by a charged sphere) from vacuum to a high dielectric solvent [142].

When considering the presence of counterions, in addition to the charge density resulting from the presence of a reference ion, an ad-

ditional charge density, ξ, is included to account for the other ions in solution as,

$$\nabla \cdot [\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) - 4\pi\xi(\mathbf{r}) \ . \tag{2.15}$$

Debye and Hückel extended the Born model by proposing an hypothesis concerning the distribution of ions around a reference ion. Their hypothesis was that, in the presence of mobile ions, the local ion concentration is given by the local electrostatic potential [141], and Equation 2.15 becomes the Poisson-Boltzmann equation which, after linearization, becomes the linear Poisson-Boltzmann equation (LPBE) [135, 141, 146]:

$$\nabla \cdot [\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] - \varepsilon(\mathbf{r})\kappa^2(\mathbf{r})\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \tag{2.16}$$

where κ^2 is the Debye length, defined as:

$$\kappa(\mathbf{r}) = \begin{cases} \left(\frac{8\pi e^2 I}{\varepsilon_{out} k_B T}\right)^{1/2} & \text{if } \mathbf{r} \text{ is in the counterion region} \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

where e is the proton charge, I is the ionic strength, T the temperature and k_B the Boltzmann constant. I is defined as a sum over all ionic species, $I = \frac{1}{2}\sum_j c_j z_j^2$; where c_j is their bulk concentrations. The Debye length is normally taken as a length scale below which mobile ions experience the solute's electrostatic potential and interact with it [141].

The electrostatic potential obtained from the LPBE equation can be used to determine several electrostatic properties. The most relevant for the present discussion being the *electrostatic energy* [10], which for a protein with a set of point charges $(q_1, q_2, ..., q_M)$ placed at positions $(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_M)$ is given by [140]:

$$W = \frac{1}{2}\sum_i q_i \phi(\mathbf{r}_i) \ . \tag{2.18}$$

The free energy difference between an initial and a final charge-state of a protein can then be defined as the difference between the W of each

[10] The *electrostatic energy* can be seen as the reversible work required to bring all charges from zero to their actual values [140].

state.

Unfortunately, for molecules with complex geometries the LPBE cannot be solved analytically. The usual approach to estimate the electrostatic potential at different positions, is to model the system as a cubic-lattice and applying a standard (numerical) finite difference procedure. In that method, the atomic partial charge values are mapped onto a grid in such a way that, the grid-points that are closer to the actual position of the charge, proportionally retain a higher fraction of it.

2.4.3 Protein titration

The titration of a protein molecule can be described in terms of the Gibbs free energy of all different ionization reactions of the fully neutral form. The objective thus becomes to express the *protonation free energy* in terms of CE/PB energy calculations.

The direct determination of reaction free energies is not always straightforward, for instance, the (de)protonation of a titrable protein group in the protein environment. The common procedure is to relate the titration of the protein with the titration of its corresponding model compounds [127, 135, 139, 147].

The protonation of a titrable site in a protein can be described using a thermodynamic cycle like the one shown in Figure 2.2 [147]. A^{sol} and A^{P} represent, respectively, the deprotonated forms of the group in the solvent (given by a model compound) and protein environments; while AH^{sol} and AH^{P} are the corresponding protonated forms. For proteins with multiple titrable sites an equivalent thermodynamic cycle must be considered for each site.

The standard free energy difference of protonating a group in the protein ($\Delta G^{\circ}_{P}(A \rightarrow AH)$) can be obtained from the other free energy terms in the thermodynamic cycle as:

$$\begin{aligned}\Delta G^{\circ}_{P}(A \rightarrow AH) &= \Delta G^{\circ}_{sol}(A \rightarrow AH) + \Delta G^{\circ}_{sol\rightarrow P}(AH) - \Delta G^{\circ}_{sol\rightarrow P}(A) \\ &= \Delta G^{\circ}_{sol}(A \rightarrow AH) + \Delta\Delta G^{\circ}_{sol\rightarrow P}(AH)\ .\end{aligned} \tag{2.19}$$

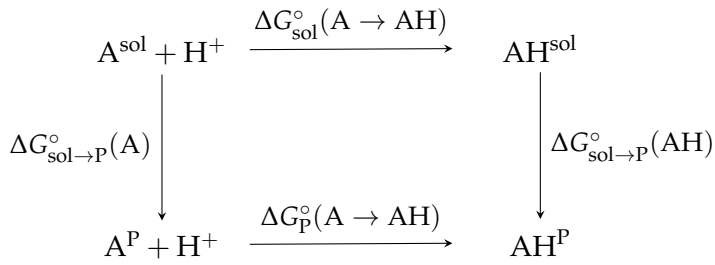


Figure 2.2: Thermodynamic cycle involving a protein and model compounds.

The free energy $\Delta G^\circ_{\text{sol}}(\text{A} \rightarrow \text{AH})$ is given by the pK_a values of model compounds (pK_mod), and $\Delta\Delta G^\circ_{\text{sol}\rightarrow\text{P}}(\text{AH})$ can be computed through CE energy differences resulting from changing from the solvent to the protein environment [127, 139, 147]. The free energy terms related with solvation and interactions between the different charges (titrating and background) can be obtained from the electrostatic potential at the sites, which can be directly estimated using the LPBE (Equation 2.16).

The charged state of all the titrable sites, i, in a protein can be defined in terms of a vector, $\mathbf{a} = (a_1, a_2, ..., a_i)$,[11] with

$$a_i = \begin{cases} 0 \text{ if site } i \text{ is neutral,} \\ 1 \text{ if site } i \text{ is charged.} \end{cases}$$

The previously defined free energy differences (Equation 2.19) can be expressed in terms of this vector. In that case the free energy $\Delta G^\circ_{\text{sol}}(\mathbf{a})$ is given in terms of the ionization free energies of the model compounds, as

$$\Delta G^\circ_{\text{sol}}(\mathbf{a}) = \sum_i a_i \Delta G_{mod,i} = -2.3 k_B T \sum_i a_i \gamma_i \text{p}K_{\text{mod},i} \ , \tag{2.20}$$

[11] While histidine is the only proteinogenic aminoacid with true tautomeric forms (in the neutral form the proton can be bound either to the N^δ or the N^ε), other titrable amino acids may have non-equivalent chemically protonable positions (proton isomers). In order to increase the realism of the CE/PB approach it is common that a_i can take as many values as the number of possible isomers plus one (the charged state). The number of possible isomers will depend on the amino acid in question, and slightly different mathematical expressions are used to estimate the protonation free energy (details in reference [148])

where γ_i is the charge (in protonic units, ± 1) of the ionized form of site i.[12]

The derivation of the expression for $\Delta\Delta G^{\circ}_{\mathrm{sol}\rightarrow \mathrm{P}}(\mathbf{a})$ is more complex and not presented; the reader can find it in reference [139]. Nevertheless, the important conclusion for the present discussion is that the protonation free energy of changing from a reference neutral state (all entries of the vector equal to zero) to another particular charge-state **a** can be written as [127, 139]:

$$\Delta G^{\circ}_{\mathrm{P}}(\mathbf{a}) = -2.3 k_B T \sum_i a_i \gamma_i \mathrm{p}K_{\mathrm{int},i} + 2 \sum_i \sum_{j<i} a_i a_j \Delta W_{ij} \ , \tag{2.21}$$

where ΔW_{ij} is essentially the interaction free energy between ionized sites i and j. And the pK_{a} value of a titrable site i in an otherwise neutral protein, pK_{int} is given by:

$$\mathrm{p}K_{\mathrm{int},i} = \mathrm{p}K_{\mathrm{mod},i} - \frac{\gamma_i}{2.3 k_B T} \Delta\Delta W_i^{env} \ , \tag{2.22}$$

where $\Delta\Delta W_i^{env} = \Delta W_i^{env}(1) - \Delta W_i^{env}(0)$, with $\Delta W_i^{env}(a_i)$ corresponding to the CE energy change due to the mapping of the site in state 0 or 1 into the protein environment (see reference [139] for further details).

Once the pK_{int} and ΔW_{ij} values are obtained one can compute the the probabilities of ionization states, as

$$p(\mathbf{a}) = \frac{\exp\left[-2.3 z(\mathbf{a}) \mathrm{pH} - \frac{\Delta G^{\circ}(\mathbf{a})}{k_B T}\right]}{\sum_{\mathbf{a'}} \exp\left[-2.3 z(\mathbf{a'}) \mathrm{pH} - \frac{\Delta G^{\circ}(\mathbf{a'})}{k_B T}\right]} \tag{2.23}$$

where $z(\mathbf{a})$ is the net charge of state **a** (i.e. $z(\mathbf{a}) = \sum_i a_i \gamma_i$).

However, if the number of titrable groups in the protein is very large, it is not possible to calculate $p(\mathbf{a})$ explicitly for all possible **a**.[13] In this case, the probability distribution can be estimated in a stochastic manner, using a Monte Carlo method as the one described in previous sec-

[12] The number 2.3 stands for ln 10.

[13] A protein with n titrable sites will have 2^n possible protonation states.

tions. We start by defining an "energy" [126, 127],

$$\mathcal{E}(\mathbf{a}) = -2.3k_BT\sum_i a_i\gamma_i[\mathrm{p}K_{\mathrm{int},i} - \mathrm{pH}] + 2\sum_i\sum_{j<i} a_i a_j \Delta W_{ij}\,, \tag{2.24}$$

then $p(\mathbf{a})$ can be rewritten as:

$$p(\mathbf{a}) = \frac{\exp\left[-\frac{\mathcal{E}(\mathbf{a})}{k_BT}\right]}{\sum_{\mathbf{a}'}\exp\left[-\frac{\mathcal{E}(\mathbf{a}')}{k_BT}\right]}\,. \tag{2.25}$$

Formally, sampling of protonation states is now identical to a MC sampling from a canonical ensemble. The Metropolis criterion, based on $\mathcal{E}(\mathbf{a})$, is used directly to accept/reject a trial state (in analogy with a trial move). In essence, this approach combines the CE-LPBE perspective on titration with the MC sampling of protonation states and is, therefore, commonly referred to as a Poisson-Boltzmann/Monte Carlo (PB/MC) method.

Once the probability distributions of **a** at different pH values are estimated, it becomes possible to calculate the titration curve of the whole protein as well as the titration curves of individual sites [127, 143]. The $\mathrm{p}K_\mathrm{a}$ of a site can then be estimated from its titration curve by considering, for instance, the pH of half-titration (i.e. $\mathrm{p}K_\mathrm{half}$).

2.5 Constant-pH MD

One of the most interesting aspects of molecular charge states is their sensitivity to environmental effects, namely pH. Unfortunately, as mentioned in the above section, CE methods seem adequate only to describe non-flexible molecules. Although, as mentioned in the earlier sections of this chapter, conformational freedom is of paramount importance to describe biomolecular systems, and hence the utility of MM/MD and MM/MC methods.

The usual way of dealing with pH in MD simulations consists in assigning fixed protonation states to the protonable groups. Those states should be coherent with the intended pH and are normally selected

based on tables of solution pK_as for the residues or on (preliminary) CE calculation with a single initial structure.[14] This has also been the standard in simulations of dendrimers.

The problems arise when multiple protonable groups are in close proximity (chances of severe cooperative effects), or highly internalized (lack of stabilizing solvation interactions) or most frequently when the pK_a of a group is near the pH intended for the simulation, which could entail that different protonation states coexist with significant probabilities [135].

Fortunately, PB/MC and MM/MD methodologies exhibit some complementarity. MM-based methods capture the conformational freedom of a single charged form of a protein and CE methods can address the non-flexible protein titration.

The idea of simultaneously exploring this complementarity is the base of some *constant-pH* MD (CpHMD) methods [128, 135]. In these methods pH is regarded as an external thermodynamic parameter.

Several constant-pH MD methods are available [149–154]. Here we will focus on the *stochastic titration* CpHMD method developed (and continuously improved) by A. Baptista, M. Machuqueiro and coworkers [128, 139, 155–158].

As depicted in Figure 2.3 the algorithm behind the stochastic titration method is a cycle with three main steps:

1. For a specific protein configuration a PB/MC step, at a given pH value, is performed to compute the protonation free energies of previously identified titrable sites and assign new protonation states for the sites according to the last MC move;

2. A solvent relaxation step follows. It consists of a short MM/MD simulation where the solute molecule is kept rigid while solvent molecules are allowed to adapt to new charge configuration. This avoids the occurrence of unfavorable solute-solvent interactions due to the change in protonation states.

[14] For this latter approach to have any meaning the initial structure should be an experimentally determined one; which, to date, excluded its application in computational studies of dendrimers.

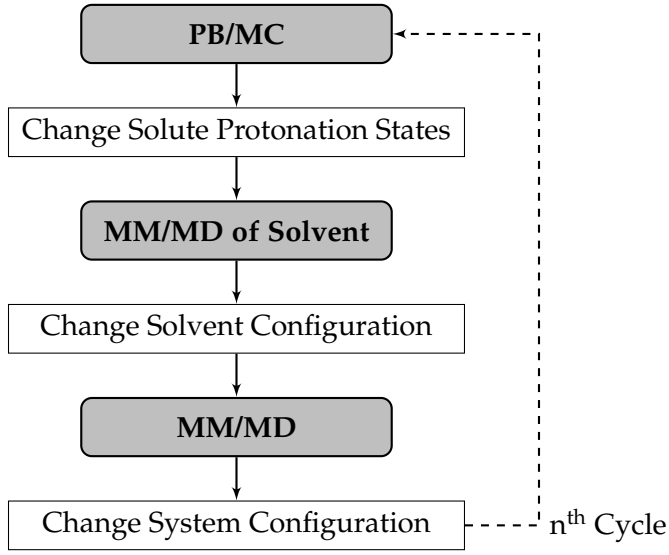


Figure 2.3: Schematic representation of the stochastic titration contant-pH MD algorithm [126, 128].

3. The productive MM/MD step. Another MM/MD simulation is performed, without any restrains on the atoms, in order to sample the configurational space of the new charge configuration. The last conformation obtained from the simulation trajectory is used as input for the following PB/MC calculation; which corresponds to a new three-step cycle.

In addition to the three-step procedure described, a reduced titration approach, similar to the one described by [155, 159], can be included.

Statistically, it can be said that the *stochastic titration* CpHMD method properly models the protonation–conformation coupling in titrable molecules, allowing also for an accurate description of the protonation equilibria that takes place at different pH values.

Chapter 3

Unraveling the Conformational Determinants of Peptide Dendrimers Using Molecular Dynamics Simulations

Luís C. S. Filipe[1], Miguel Machuqueiro, Tamis Darbre and António M. Baptista

Contents

[1] In this work I performed all the simulations and analysis.

3.1 Summary

Peptide dendrimers are synthetic tree-like molecules composed of amino acids. There are at least two kinds of preferential structural behaviors exhibited by these molecules, which acquire either compact or noncompact shapes. However, the key structural determinants of such behaviors remained, until now, unstudied. Herein, we conduct a comprehensive investigation of the structural determinants of peptide dendrimers by employing long molecular dynamics simulations to characterize an extended set of third generation dendrimers. Our results clearly show that a trade-off between electrostatic effects and hydrogen bond formation controls structure acquisition in these systems. Moreover, by selectively changing the dendrimers charge we are able to manipulate the exhibited compactness. In contrast, the length of branching residues does not seem to be a major structural determinant. Our results are in accordance with the most recent experimental evidence and shed some light on the key molecular level interactions controlling structure acquisition in these systems. Thus, the results presented constitute valuable insights that can contribute to the development of truly tailor-made dendritic systems.

3.2 Introduction

Dendrimers are a family of highly branched compounds that share a common layout where wedges emerge radially from a core by means of a regular branching pattern [7, 9, 10, 13, 93, 160]. The tree-like architecture exhibited by these molecules grants them a characteristic multivalency as well as distinctive internal microenvironments [10, 16]. Topologically, dendrimers are characterized by three distinct regions: core, branches, and periphery, with the latter being typically composed of functionalized end-groups [7, 16, 22].

Peptide dendrimers are a specific kind of dendrimers formed by alternating proteinogenic amino acids with branching diamino acids such as lysine [10, 36, 37, 40–42, 161, 162]. Because they are composed of the same constituents as proteins, peptide dendrimers demonstrate

higher biocompatibility and biodegradability than some other common synthetic agents [17, 36, 40, 48, 163–165]. Therefore, it comes as no surprise that, over the past years, these promising molecules have already rendered several synthetic models with potential technological and biomedical relevance [40, 42, 44, 97], namely, catalytic peptide dendrimers able to perform ester hydrolysis or aldol-type reactions [47, 50, 54, 58–60, 166], metal binding peptide dendrimers [55, 64, 65], multivalent lectin binding systems [67–70, 167], drug delivery agents [62, 63], or antimicrobial peptide dendrimers [72].

Although the synthesis of these systems is becoming increasingly straightforward, two main issues remain mostly unaddressed and hindering their widespread application: (1) little is known about these molecules structure in solution and the influence played by each of its constituents on the overall fold, and (2) the selection of which amino acids to use for specific functions is ruled by chemical reasoning and trial and error screening of large combinatorial libraries.

We have recently used long molecular mechanics/molecular dynamics (MM/MD) simulations to investigate the conformational preferences of third-generation peptide dendrimers know to act as synthetic models for cobalamin-binding proteins [100]. That study highlighted the extreme flexibility displayed by peptide dendrimers, showing a myriad of conformational states accessible to them without a clear "folded" state. This conformational plasticity is probably the reason why these molecules have not yielded to structural characterization using common experimental techniques, and why computational methods have become a pervasive tool to explore the features of dendrimers at the molecular level [69, 70, 95, 97, 98, 168].

Despite the high flexibility that seems to characterize peptide dendrimers, on the aforementioned MM/MD study two markedly distinct conformational behaviors were observed, with some dendrimers favoring mainly noncompact (loose) conformations (e.g., B1 in Table 3.1), while others prefer more compact configurations (e.g., C1 in Table 3.1). Moreover, the conformational differences between noncompact/compact dendrimers were adequately captured by two-dimensional energy

Table 3.1: Residue composition and total charge for each of the peptide dendrimers studied.

	residues at each position[a]											
dendrimer	X_1	X_2	B_1	X_3	X_4	B_2	X_5	X_6	B_3	X_7	X_8	charge (e)
Previously Studied												
B1[b]	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	Glu	Dap	Ser	Glu	-13
C1[b]	Ala	Arg	Dap	Thr	His	Dap	Tyr	Glu	Dap	Gly	Ser	-3
Branching Residues Effect												
B1-series	Asp	Cys		Tyr	Amb		Ala	Glu		Ser	Glu	
B1Lys3			Lys			Lys			Lys			-13
B1[b]			Dap			Lys			Dap			-13
B1Orn			Dap			Orn			Dap			-13
B1Dab			Dap			Dab			Dap			-13
B1Dap[c]			Dap			Dap			Dap			-13
C1-series	Ala	Arg	Dap	Thr	His		Tyr	Glu	Dap	Gly	Ser	
C1Lys						Lys						-3
C1Orn						Orn						-3
C1Dab						Dab						-3
C1[b]						Dap						-3
Charged Residues Effect												
NE-series	Asp	Cys	Dap	Tyr	Amb	Dap	Ala		Dap	Ser		
NE12[c]								Glu			Glu	-13
NE8								Gln			Glu	-9
NE4								Glu			Gln	-5
NE0								Gln			Gln	-1
CE-series	Ala	Arg	Dap	Thr	His	Dap	Tyr		Dap		Ser	
CE12								Glu		Glu		-11
CE8								Gln		Glu		-7
CE4								Glu		Gln		-3
CE0								Gln		Gln		+1

[a]Residue positions in accordance with Figure 3.1. For each dendrimer series, the common amino acid residues are shown in its first row (B1-series, C1-series, NE-series, and CE-series), while only the variable ones are shown in the subsequent sequence-specific rows. Standard three-letter abbreviations are used for proteinogenic amino acids, Amb for 4-aminomethyl(benzoic) acid, Orn for Ornithine, Dab for *L*-2,4-diaminobutanoic acid and Dap for *L*-2,3-diaminopropanoic acid. [b]The simulation data for B1 and C1 was taken from ref [100]. [c]Note that B1Dap and NE12 are the same dendrimer.

landscapes, where noncompact-type dendrimers present a clear downhill propensity whereas the compact-type ones exhibited rough landscapes [100].

The reasons for these different conformational behaviors remained at the time unstudied, but, based on the amino acid composition of the dendrimers and the experimental data available [55, 64] it was proposed

that the placement of charged residues and/or the length of the branching residues could act as key structural determinants.

The significance of charge distribution has been the focus of recent atomistic and coarse-grained molecular simulation studies in polyamidoamine (PAMAM) dendrimers, revealing that the degree of amine (de)protonation is responsible for dramatic structural transitions. The outcome of those studies provided a clearer depiction of PAMAM dendrimers shape, swelling and loading ability [83, 169]. In fact, computational methods have become a pervasive tool to explore and characterize the effect of electrostatics [90, 93, 170, 171] and branching symmetry [101] in different dendritic systems.

Herein, we perform a comprehensive investigation on the structural determinants of peptide dendrimers by sampling the conformational preferences of several series of third-generation peptide dendrimers (in a total of 17 distinct peptide dendrimers) using multiple and long MM/MD simulations (see Table 3.1). All the dendrimers studied here have the same topology (Figure 3.1).

We started by investigating the role played by the length of the branching residues side-chains. For that we simulated two series of dendrimers with sequentially shorter/longer branching residues side-chains. We used as starting points the sequences of B1 (noncompact behavior) and C1 (compact behavior) and systematically replaced their branching residues with: lysine [side-chain: $(CH_2)_4NH_2$], ornithine [side-chain: $(CH_2)_3NH_2$], diaminobutanoic acid [side-chain: $(CH_2)_2NH_2$] or diaminopropanoic acid [side-chain: CH_2NH_2]. See B1- and C1-series in Figure 3.1 and Table 3.1 and section 3.4.1 for details.

A recent experimental study by Reymond and coworkers [64] suggests that the size of peptide dendrimers increases with the number of negatively charged residues, an hypothesis also consistent with our previous study [100]. To investigate the conformational role played by negatively charged residues, we simulated peptide dendrimers with different combinations of negatively charged glutamic acid and neutral glutamine residues in the second and third dendrimer generations (see NE- and CE-series in Figure 3.1 and Table 3.1, and also section 3.4.2 for de-

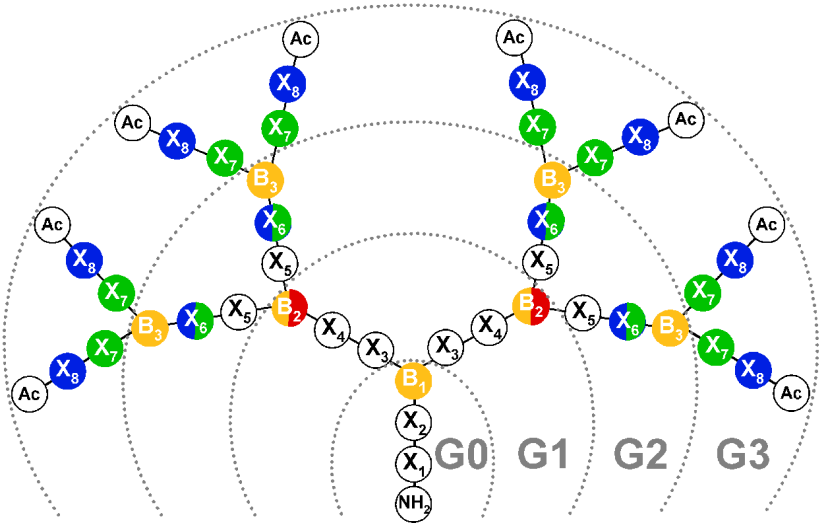


Figure 3.1: Schematic topology of the peptide dendrimers investigated. The X_n and B_n residues of each dendrimer are indicated in Table 3.1. The amino acid residues *modified* in each of the dendrimer series studied are highlighted using a color code: B1-series (yellow), C1-series (red), NE-series (blue), and CE-series (green). Also highlighted are the dendrimer generations (G0–G3). Ac = acetyl group (N-terminal cap). NH_2 = amine group (C-terminal cap).

tails). Three of the dendrimers studied in the present work (NE12, NE4, and NE0) have been previously synthesized and studied by Reymond and coworkers [64].

To the best of our knowledge, this is the first comprehensive computational study of the sequence-structure relationship in peptide dendrimers using several systematic series of molecules. The results and conclusions presented here might, in the near future, help to rationalized some of the available experimental data, contributing to the development of truly tailor-made peptide dendrimers.

3.3 Computational Methods

3.3.1 Simulation Details

The simulation protocol employed in the present work is identical to the one used in ref [100]. Therefore, we mention here only its key features. Note that the data presented for dendrimers B1 and C1 results from simulations already published in that work.

We have used PyMOL [172] (version 0.99rc6) to obtain an initial set of 3D coordinates for each peptide dendrimer in Table 3.1. All the dendrimers were simulated with an amino group (NH_2) attached to the core C-terminus, and acetyl (Ac) groups to the peripherical N-termini.

The GROMACS package [173–175], version 4.0.2, and the GROMOS 53A6 force field [114] were used to perform the MM/MD simulations. Most of the amino acid blocks considered for the topology of peptide dendrimers were already available in the GROMOS 53A6 set. The topology blocks for 4-aminomethyl(benzoic) acid (Amb), *L*-2,3-aminopropanoic acid (Dap) and the branching lysine residues have been described previously [100]. New topology blocks were constructed for *L*-ornithine (Orn) and *L*-2,4-diaminobutanoic acid (Dab) assuming the transferability of the force field (blocks provided in Appendix B, page 132).

The charges assigned to each titratable residue were the ones typically present at pH 7. The suitability of the protonation states chosen was confirmed *a posteriori* by pK_a calculations over the conformation ensembles obtained from the simulations (Appendix B, page 135).

The nonbonded interactions were treated with a twin-range cutoff of 8/14 Å and neighbor lists updated every 10 fs. The reaction-field method [117], with a relative dielectric constant of 54.0 [176], was used for the long-range electrostatic interactions. Molecular dynamics simulations were performed by integrating the equations of motion using the Verlet leapfrog algorithm with a time step of 2 fs, and the system coordinates (snapshots) were saved every 10 ps for further analysis. The temperature was kept constant at 298.15 K using the Berendsen coupling scheme [177] with a relaxation time of 0.1 ps. A Berendsen isotropic

pressure coupling [177] was used at 1 bar, with a relaxation time of 0.5 ps and an isothermal compressibility of 4.5×10^{-5} bar^{-1}.

All simulations were done with explicit solvent, using single point charge (SPC) [178] water molecules in rhombic dodecahedral boxes, while applying periodic boundary conditions. The final systems contained about 32 to 50 thousand atoms.

All structures were subject to energy minimization and initialization procedures previous to the production stage of the simulation. To ensure the formation of unbiased initial conformations, during some initialization steps the partial charges of all dendrimer atoms were changed from their reference values to a value of +0.1 e [100]. This change in atomic electric charges promotes the repulsion among all dendrimer atoms, originating a generic initial structure that corresponds to the most "stretched" conformation of each dendrimer. Starting with these fully extended configurations, 10 molecular dynamics simulations of 100 ns were performed for each dendrimer, accounting for a total of 1 µs per dendrimer. The replicates of each peptide dendrimer were started from the same optimized system but with different sets of random velocities. By concatenating the equilibrated trajectories of the different replicates, one obtains the total amount of production simulation.

The equilibration time of the different replicates was determined by monitoring the radius of gyration [179]. The systems showed to be equilibrated at different time lengths (ranging from 5 to 20 ns) and only the equilibrated trajectories were used for subsequent analyses, which were done using GROMACS [173–175] or in-house tools. The statistical uncertainty of the radius of gyration was computed using the jackknife method [180].

3.3.2 Energy landscapes

Macromolecules are normally quite elaborate (and dynamic) molecular systems that can adopt a plethora of geometrical configurations, often with similar energies. Energy landscapes (or surfaces) can, in principle, be used to characterize and categorize all possible conformations [181].

A molecule composed by N atoms in a 3D-space has $3N$ degrees

of freedom, of which $3N-6$ are internal ($3N-5$ if the molecule is linear); the system conformational energy landscape can be naturally expressed as an hyper-dimensional landscape including all $3N-6$ degrees of freedom [105]. Ideally, if a system's energy landscape could be fully mapped, any property or characteristic of such system could be derived from it. However, the sheer number of degrees of freedom for even simple molecules leads to a vast conformational space, making exhaustive exploration an unfeasible task.

Fortunately, the energy landscape concept is entirely generic; landscapes may correspond to potential energy terms of a system as a function of the coordinates of all its atoms, or they may also include thermal energy, thereby corresponding to the free energy of the system as a function of only some coordinates of interest [105].

The description of a landscape is made by analogy with natural landscapes. An energy landscape can be fairly smooth or corrugated with peaks and valleys of distinct magnitudes and shapes; the valleys (or basins) refer to conformational substates and the saddle points correspond to minimum energy barriers [181]. The depth of a basin is its enthalpy, while the width near this local minimum reflects the entropy of that state.[2]

To characterize the conformational space of peptide dendrimers, we must identify the accessible energy basins and minima on the energy landscape.

To achieve this, we have determined probability density functions, $P(\mathbf{r})$, in a two-dimensional (2D) representation space using as structural coordinates the radius of gyration (R_g) [179],

$$R_g = \left(\frac{\sum_i |\mathbf{r}_i^2| m_i}{\sum_i m_i} \right)^{1/2} , \tag{3.1}$$

and the root mean-square deviation (rmsd) [182] to the central structure

[2]The entropy is a measure of the number of thermally accessible states, so a wide/shallow basin corresponds to greater entropy than does a narrow/deep basin (entropy/enthalpy compensation) [105].

[100, 183]:

$$\mathrm{rmsd}(A,B) = \left(\frac{1}{\sum_{i=1}^{N} m_i} \sum_{i=1}^{N} m_i |\mathbf{r}_i(A) - \mathbf{r}_i(B)|^2 \right)^{1/2} . \qquad (3.2)$$

The radius of gyration provides a measure of a dendrimer compactness [88, 171, 184–188], while the rmsd quantifies the (dis)similarity among pairs of conformations. The rmsd to a "central" structure [183] was computed using all dendrimer atoms, according to a combinatorial procedure accounting for quasi-symmetry described elsewhere [100]. The R_g-rmsd representation space was built using snapshots from the final concatenated trajectories at intervals of 0.2 ns. The probability density functions were estimated using a Gaussian kernel estimator [189], while employing grids of $(0.002\,\text{Å})^3$.

Energy surfaces were computed from $P(\mathbf{r})$ according to

$$E(\mathbf{r}) = -RT\, ln \frac{P(\mathbf{r})}{P_{max}} , \qquad (3.3)$$

where $P(\mathbf{r})$, P_{max}, T, and R are, respectively, the probability density function, its maximum, the absolute temperature, and the ideal gas constant [100, 183]. For simplicity, we will hereafter refer to $E(\mathbf{r})$ as "energy", although it, in fact, represents a conditional free energy [181, 183, 190, 191]. This quantity has indeed a free energy character, although that does not follow alone from its logarithmic dependence on a probability, but rather from the fact that such probability refers to a set of microstates. Strictly speaking, this quantity is a partial free energy (or conditional free energy or potential of mean force), since the corresponding sum-over-states is over all degrees of freedom except the structural coordinates explicitly considered (e.g. R_g and rmsd).

We group the configurations expressed in the 2D conformational space using as clustering condition the confinement within a common energy basin, while considering different energy cutoffs [191].

3.4 Results and Discussion

3.4.1 Conformational Relevance of Branching Residues

The clearly distinct conformational preferences observed in ref [100] for B1 and C1 dendrimers result from differences in the composition of these two molecules. One such difference is the length of the side-chain of the B_2 branching residues (see Table 3.1). While those positions are occupied in C1 by Dap residues with short side-chains [-CH_2NH_2], in B1 they consist of Lys residues with much longer and freely rotating side-chains [-$(CH_2)_4NH_2$] that may confer greater flexibility to the dendrimer.

To investigate the importance of the length in the branching residues side-chains, we conducted simulations on two series of peptide dendrimers, the B1- and C1-series (Table 3.1). The two dendrimer series are very different in their amino acid sequences, with the B1-series being based on the noncompact dendrimer B1 while the C1-series is based on the compact C1. The sole distinction between the dendrimers in each series is the branching residue in the second generation (B_2 in Figure 3.1), where we sequentially replace the branching residue by residues with longer/shorter side-chains, namely: Lys [-$(CH_2)_4NH_2$], Orn [-$(CH_2)_3NH_2$], Dab [-$(CH_2)_2NH_2$], or Dap [-CH_2NH_2]. As a consistency test we have also simulated B1Lys3 (Table 3.1), a peptide dendrimer similar to B1 with lysines at all branching residues. With this dendrimer we expect to investigate if branching residues occupying different positions in the overall topology have a similar contribution to structure acquisition.

The R_g probability density histograms for the different dendrimers are displayed in Figure 3.2. As can be observed, although there are some differences in the R_g profiles of dendrimers within the same series, changing the branching residues has little influence on their overall compactness. This seems to be true for dendrimers known to have mostly compact (C1-series) or noncompact (B1-series) structures.

The fact that B1Lys3 has a R_g probability density distribution similar to the other B1-series dendrimers highlights the fact that the conforma-

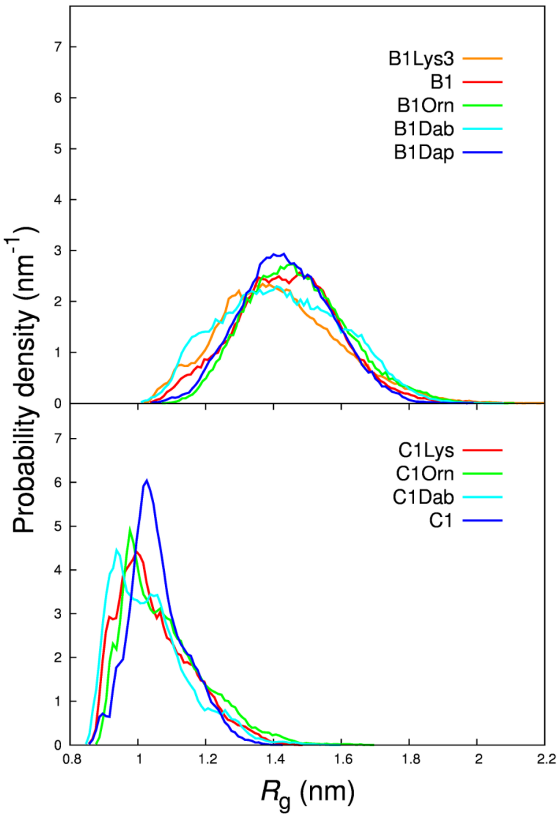


Figure 3.2: Radius of gyration probability density histograms for the B1- and C1- dendrimer series (top and bottom respectively).

tional preferences of these systems seem also to be independent of the type and position of branching unit used. These conclusions are further supported by the dendrimers average radius of gyration values (Table 3.2).

It must be mentioned that the maximum R_g value observed for B1Lys3 (2.24 nm) is higher than the maximum value observed for the other B1-series dendrimers (approximately 2.05 nm). Therefore, replacing all branching positions in the dendrimer by the branching residue with the longest side-chain studied here, lysine, indeed promotes the existence of more extended conformational states, though those states are seldom populated.

A detailed analysis of 2D energy landscapes employing R_g and rmsd

Table 3.2: Average Radius of Gyration for each B1- and C1-Series Peptide Dendrimer.

dendrimer	$\langle R_g \rangle$ (nm)
B1-Series	
B1Lys3	1.42±0.06
B1[a]	1.44±0.04
B1Orn	1.47±0.03
B1Dab	1.43±0.06
B1Dap	1.44±0.02
C1-Series	
C1Lys	1.05±0.04
C1Orn	1.08±0.04
C1Dab	1.04±0.04
C1[a]	1.06±0.03

[a]The simulation data for B1 and C1 was taken from ref [100].

as structural coordinates (Appendix B, page 137) shows that peptide dendrimers of the same series have extremely similar landscapes. In particular, the dendrimers from the B1-series have smooth landscapes with a clear downhill propensity, where no significant conformational clusters are identifiable and the energy minima correspond to noncompact conformations. In contrast, dendrimers from the C1-series have rough landscapes with some conformational clusters, and energy minima that correspond mainly to compact dendrimer conformations.

Previous studies [98, 100] suggested that peptide dendrimers have some conformational robustness to changes in sequence. This idea is also supported by the data presented here.

3.4.2 Conformational Relevance of Charged Residues

Recent experimental evidence suggest that the size of peptide dendrimers increases with the number of negatively charged residues [64]. To investigate this effect and its causes, we have built and simulated peptide dendrimers with different combinations of negatively charged glutamates and neutral glutamines in the second (G2) and third (G3)

generation layers (NE- and CE-series in Table 3.1).

Again, we used two different dendrimer series: the NE-series, derived from the noncompact NE12 dendrimer (NE12 is the same as B1Dap); and the CE-series, derived from the compact C1 dendrimer after a Gly $\rightarrow$ Gln mutation at position X_7, for consistency of the tests. The Gln/Glu replacements affect 4 residues in the G2 shell (e.g., position X_6 in Figure 3.1), 8 in the G3 shell (e.g., position X_8 in Figure 3.1), and 12 when performed in both G2 and G3. While the G2 replacements are done at position X_6 in both series, the G3 replacements are done at position X_8 for the NE series but at position X_7 for the CE series, in order to preserve the Ser residues present at G3 in all the series investigated here.

In addition to the CE- and NE-series, we have also performed simulations in a BE-series obtained from Gln/Glu replacements in B1, thus differing from the NE-series in having Lys instead of Dap at position B2. Since the results obtained for NE and BE are mostly identical, we give the emphasis to the former, for which experimental data is available (see below); the results for BE are shown in Appendix B (page 138 to 140).

Dendrimer Compactness

Figure 3.3 presents the R_g probability density histograms for the different NE- and CE-series dendrimers, with the average values shown in Table 3.3. The data clearly shows that charge has a major influence on the conformational preferences of peptide dendrimers, being directly related with compactness: the more a dendrimer is negatively charged, the less compact its structures are. This is not entirely surprising, and a parallelism can even be established with the crucial role played by electrostatics in protein folding. What is somewhat unexpected is the essentially linear correlation between the number of negatively charged residues and compactness, as seen in Figure 3.4. For the BE-series see Appendix B, page 139. Simulations of the CE8 and CE4 dendrimers treating long-range electrostatics using a particle mesh Ewald method, instead of a reaction field, show a similar trend (see Appendix B, page 141).

Table 3.3 compares the computed average R_g with the hydrodynamic

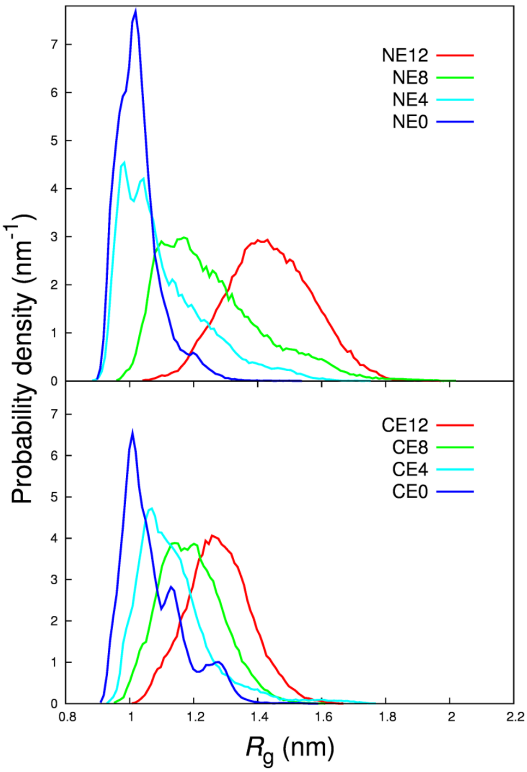


Figure 3.3: Radius of gyration probability density histograms for the NE- and CE-series (top and bottom respectively).

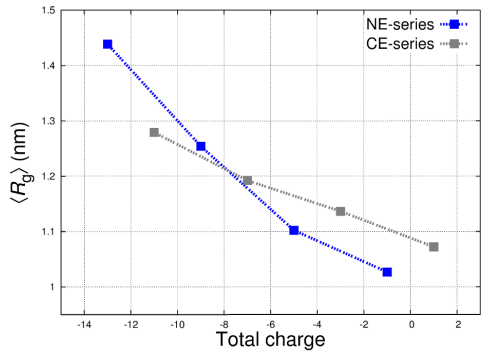


Figure 3.4: Dendrimer total charge versus average radius of gyration for the NE- and CE-series.

radius experimentally measured [64] using diffusion NMR for the NE-series dendrimers; although different, the two properties are usually similar or approximately proportional [101, 192, 193]. The simulation results are consistent with the experimental data, with both exhibiting the same overall trend, attesting the suitability of the computational procedures employed.

Table 3.3: Computed average R_g values and experimentally measured hydrodynamic radii for the NE- and CE-Series dendrimers.

dendrimer	$\langle R_g \rangle$ **(nm)**	R_h[a] **(nm)**
	NE-Series	
NE12	1.44±0.02	1.67±0.05
NE8	1.25±0.06	n.a
NE4	1.10±0.04	1.53±0.01
NE0	1.03±0.02	1.48±0.01
	CE-Series	
CE12	1.28±0.02	n.a
CE8	1.19±0.02	n.a
CE4	1.14±0.03	n.a
CE0	1.07±0.05	n.a

[a]The experimental R_h values were obtained from ref [64]. Neither NE8 nor the CE-series were synthesized/measured in that work.

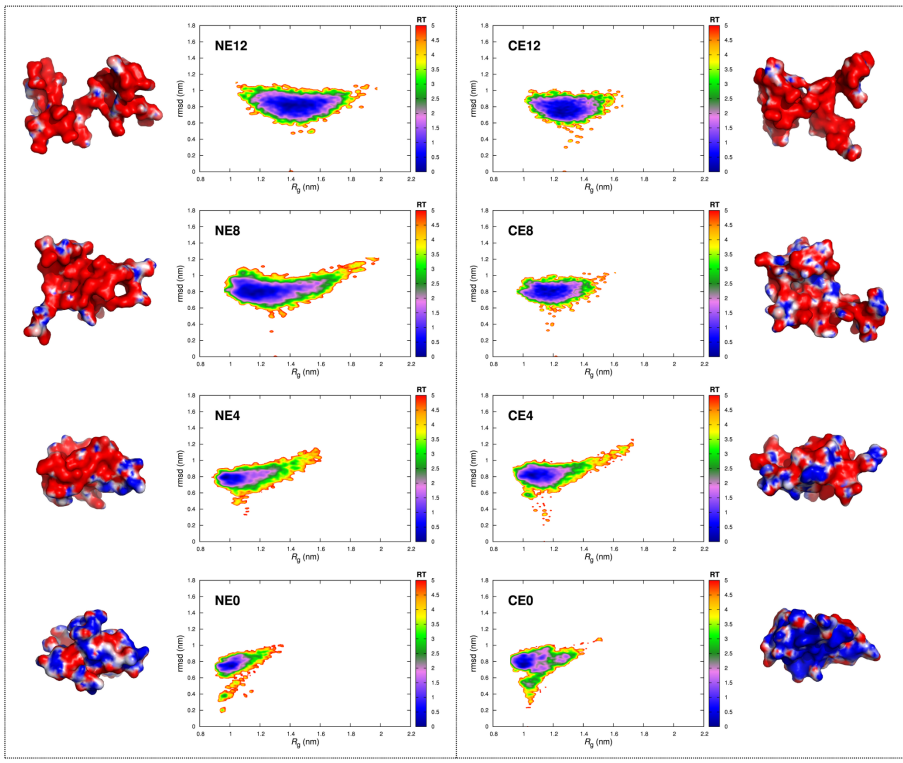


Figure 3.5: Energy profiles and corresponding lowest energy conformers for the NE- (left) and CE-series (right). The energy profiles use R_g and rmsd as structural coordinates. The lowest energy conformers are drawn using electrostatic potential surfaces computed with MEAD [194].

Energy Landscapes

Performing Glu/Gln residue replacements in a peptide dendrimer sequence noticeably modifies its conformational preferences, markedly changing the features of its energy landscape (Figure 3.5).

The roughness of the landscapes increases with the decrease in the number of negatively charged residues. In fact, the downhill propensity that characterizes peptide dendrimers with noncompact behaviors tends to vanish with the decrease of the overall dendrimer charge, while the number of identifiable conformational clusters increases when the systems have mainly neutral residues.

The resemblance between landscapes of dendrimers from the NE-

and CE-series that have the same combination of Glu/Gln residues is remarkable, and clearly points to conformational dependence on electrostatic effects. The results presented in the Appendix B for the BE-series (page 140) also support this conclusion. Undoubtedly, electrostatics is a major determinant of the structure of peptide dendrimers and has to be taken into account when devising new dendritic systems.

Hydrogen Bond Preferences

Replacing Glu residues by Gln ones is not simply a matter of replacing charged residues by neutral ones, because Glu residues have on their side-chain a carboxyl group with its two oxygen atoms that can act only as hydrogen bond acceptors, whereas Gln residues have on their side-chains an amide group, where the oxygen can act as an acceptor while the nitrogen atom can act as donor for hydrogen bonds. These differences in hydrogen bond capability are significant when comparing, for instance, NE12 with NE0. These differences are also clear when comparing the dendrimers average radius of gyration with the average number of hydrogen bonds found during the simulations, where an approximate linear relationship is also observed (Appendix B, page 142).

To further analyze this issue, we have computed matrices where we plot all potential hydrogen bond donors against all hydrogen bond acceptors and calculate the frequency that the hydrogen bond occurs along the simulations. The frequency of a hydrogen bond is calculated simply by screening each conformation in the ensemble for the presence of the hydrogen bond and normalizing the sum by the total number of conformations. Hydrogen bonds were defined using a geometrical criterion with a maximum donor-acceptor distance of 3.5 Å and a hydrogen-donor-acceptor angle inferior to 30° [195, 196]. In this way, we describe the hydrogen bond network that characterizes each of the dendrimers. The results are presented in Figure 3.6, allowing the identification of the most pervasive hydrogen bonds and to what generation(s) they belong. The matrices for the CE- and BE-series show the same overall hydrogen bond patterns as the ones present here for the NE-series and are provided in the Appendix B (page 143 and 144).

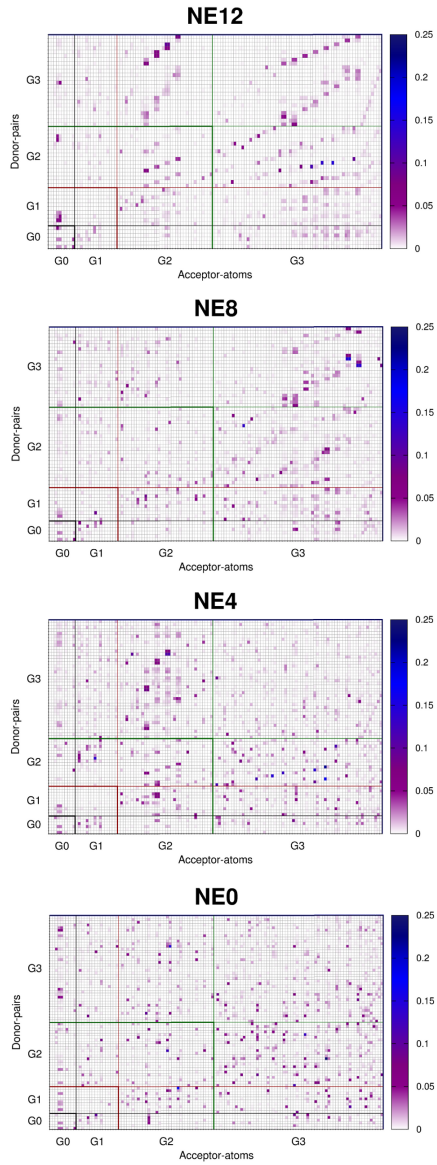


Figure 3.6: Hydrogen bond (HB) matrices for the NE-series dendrimers plotted by generation (see Figure 3.1). The matrices were obtained by plotting all potential HB donor-pairs (each donor-pair consists of a donor atom and one specific hydrogen atom bonded to it) against all HB acceptor atoms and representing the percentage of occurrence of a HB between each donor and acceptor using a color gradient.

When comparing the hydrogen bond networks of the four NE-series dendrimers it becomes clear that the incorporation of Gln residues into the dendrimers sequences spreads the hydrogen bonds formed from the third generation (case of NE12) to all over the different dendrimer generations (e.g., NE0). Nonetheless, the networks presented here are all quite transient.

Compact dendrimers, such as NE0 or CE0, have a complex and intricate hydrogen bond network that is probably responsible for stabilizing compact dendrimer conformations. Moreover, the hydrogen bond distribution patterns for these dendrimers suggests that even in compact dendrimers, there is not a clearly folded structure, but rather an heterogeneous set of possible compact conformations. As the number of donor-pairs (each donor-pair consists of a donor atom and one specific hydrogen atom bonded to it) is decreased by substituting Gln residues by Glu, together with the charge repulsion created by the presence of several Glu residues, the number and frequency of the hydrogen bonds also decreases.

The differences between the dendrimers that compose the NE-series lie solely on the side-chains of the residues present (Glu or Gln). Therefore, it is important to evaluate which atomic constituents are contributing to the increased hydrogen bond formation when the dendrimers acquire a more compact structure. Table 3.4 shows the average number of hydrogen bonds formed between donor and acceptor atoms categorized according to: main-chain atoms, atoms in the side-chain of neutral residues (Gln, among others), and atoms in the side-chains of negatively charged residues (Glu in X_8 or X_6 and one Asp residue in X_1).

Most of the trends observed in Table 3.4 can be easily understood in terms of the amounts of available residues. As the number of Gln residues increases, the neutral side-chains establish more hydrogen bonds among themselves and with the invariant main-chain, which is expected. Also, as the number of Glu residues increases, the negative side-chains establish more hydrogen bonds with the invariant main-chain, which is again not surprising. As for the hydrogen bonds between negative and neutral side-chains, they are more frequent when

Table 3.4: Average number of hydrogen bonds formed between donor-acceptor atoms which are part of Main-Chain, Side-Chains of neutral residues (neutral), and side-chains of negatively Charged Residues (Negative).

		acceptor		
dendrimer	**donors**	main-chain	neutral	negative
NE12	main-chain	**3.57**	0.67	5.19
	neutral	0.49	0.17	1.10
NE8	main-chain	**5.39**	1.57	3.95
	neutral	1.18	0.48	1.93
NE4	main-chain	**8.35**	2.81	3.07
	neutral	2.71	1.13	1.83
NE0	main-chain	**9.14**	5.03	0.57
	neutral	3.67	2.25	0.56

a significant number of both Glu and Gln residues exist (in NE8 and NE4), which is also understandable.

However, a somewhat unexpected feature is that the number of hydrogen bonds involving only the invariant main-chain steadily decreases with the addition of negative residues, a trend that must result from an indirect effect of the side-chains. Inspection of the table suggests that this may partly be seen as a sequestration of main-chain donors by negative side-chains, although this is an oversimplifying view of the complex and transient interplay between the different types of donors and acceptors. In any case, given the fundamental role of main-chain hydrogen bonds in maintaining protein structure, it is not surprising that they can be also a determinant for the compactness of peptide dendrimers.

Overall, the results presented along this article point to a joint effect of electrostatics and hydrogen bonds in governing structure acquisition in peptide dendrimers.

3.5 Concluding Remarks

Dendrimers in general and peptide dendrimers in particular pose a promising and conceptually interesting synthetic approach to devise novel functional molecules. Nonetheless, a deeper understanding of the molecular level properties and interactions of these molecules is mandatory.

In this work, a series of long MM/MD simulations were performed over a large set of third generation peptide dendrimers to identify key structural determinants. In particular, we investigated the structural role played by branching and charged residues.

The type of branching residues used in peptide dendrimers (as long as they are diamino acids) seems to have little influence on the conformational preferences exhibited by these systems. In particular, the hypothesis that the decrease of compactness results from the presence of longer branching residues is not corroborated by our results. Nonetheless, such effect may be more relevant in peptide dendrimers topologically distinct from the ones studied here (for instance with less spacer residues).

Electrostatics seems to play a crucial role in these molecules behavior. By manipulating the number and placement of charged residues in peptide dendrimer topologies, we can condition and perhaps even anticipate the conformational trends exhibited. This idea is supported mainly by changes in compactness and energy profiles of the dendrimers investigated. Furthermore, the hydrogen bond network established among the dendrimer residues (especially intra main-chain bonds) cannot be neglected. A joint effect of electrostatics and hydrogen bonds seems be the defining factor for these molecules structural behavior in solution. Consequently, the results presented here can be used in future works to predict more stable "folds" in new peptide dendrimers.

The present results provide also an explanation for the different conformational behaviors previously observed [100], but more importantly, they show the existence of intermediate conformational behaviors, not just compact and noncompact, thus providing a useful framework for the interpretation of the available experimental data [42, 55, 64]. Inter-

estingly, the structural behaviors observed here for peptide dendrimers with different net charges are comparable to the behaviors observed in G4 PAMAM dendrimers, where the sequential protonation (charge increase) of peripheral and inner-plus-peripheral amine groups causes a major structural rearrangement from a "dense core" (maximum density at the dendrimer core) to a "dense shell" (maximum density at the periphery) [83, 169].

This work also raised the question of what would be the conformational behavior of peptide dendrimers at pH values different from neutrality. For example, a decrease of pH could cause a corresponding neutralization of the Glu residues and a likely transition on conformational preferences. Recently Lee et al. [185] used molecular dynamics simulations to investigate the effect of pH on G4 PAMAM dendrimers grafted with arginine and histidine residues. They showed that the neutralization of histidine residues at neutral pH induces a conformational transition from a "dense shell" (pH 5) to a "dense core" configuration (pH 7), highlighting the structural importance of titratable amino acid residues in dendritic molecules. Similar pH-induced structural transitions are conceivable in the framework of peptide dendrimers. Studies are currently being conducted to elucidate this issue.

Overall, the conclusions obtained in this work are a contribution to the detailed understanding of dendritic systems composed of amino acids, and a step toward the establishment of a solid basis for the rational development of novel systems.

3.6 Acknowledgments

We thank Jean-Louis Reymond and Sara R. R. Campos for helpful and constructive discussions. We acknowledge the financial support from Fundação para a Ciência e a Tecnologia, Portugal, through project grants PTDC/QUI-QUI/100416/2008, PEst-OE/QUI/UI0612/2013, and PEst-OE/EQB/LA0004/2011, and doctoral fellowship SFRH/BD/76085/2011. We also acknowledge the support from EU COST Action CM1102.

Chapter 4

Exploring the structural properties of positively charged peptide dendrimers

Luís C. S. Filipe[1], Miguel Machuqueiro, Tamis Darbre and António M. Baptista

Contents

[1]In this work I performed all the syntheses, simulations and analysis.

4.1 Summary

We report a combined experimental and computational approach to study the structural behavior of positively charged peptide dendrimers. Third-generation dendrimers containing combinations of positive/neutral amino acid residues in the different dendrimer generations were synthesized and their compactness evaluated using diffusion NMR. Molecular dynamics simulations were performed to obtain a comprehensive description of the molecular-level phenomena substantiating the structural differences observed. Comparison of the results presented with previous findings reveals a striking charge-dependent tendency in these systems, where the simple number and placement of charged amino acids in the sequence allows an extensive control over the exhibited structural features.

4.2 Introduction

Dendrimers are a family of ramified synthetic molecules where wedges emerge radially from a core by means of a regular branching pattern [7, 16]. Peptide dendrimers are a class of dendrimers formed by alternating spacing amino acids with branching diaminoacids such as lysine (see Figure 4.1 for an example) [37, 42]. These peptidic tree-like molecules have attracted considerable interest due to their inherent multivalency and the possibility of constructing molecules with a variable number and type of amino acids within the dendrimer branches and core. This topologic versatility along with the possibility of grafting different functional groups to the dendrimers end-groups has sparked the development of several applications ranging from catalytic peptide dendrimers [47] and metalloprotein models [55, 64] to antimicrobial [74] and drug delivery agents [63].

The efficiency of these applications is deeply coupled to the structural features exhibited by peptide dendrimers in solution. This has been exemplified by sequence-activity studies where the efficacy of different active dendrimers was quantified and compared to their overall compactness in solution; which was evaluated based on the den-

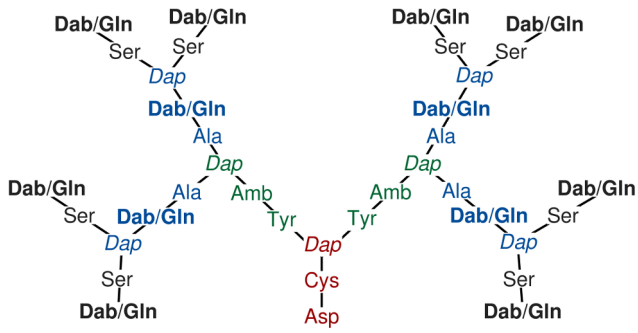


Figure 4.1: Dendrimers topology and sequence. The residues are colored according to their generation: G0 (red), G1 (green), G2 (blue) and G3 (black). *Dap*: *L*-2,3-diaminopropionic acid (branching residue); Amb: 4-aminomethyl(benzoic) acid; **Dab**: *L*-2,4-diaminobutyric acid.

drimers hydrodynamic radius (R_h). For instance, in membrane disrupting antimicrobial peptide dendrimers [72, 74], as well as in models for cobalamin-binding proteins [55, 64, 102], the more efficient dendrimers are consistently more compact than less efficient ones.

Understanding how a dendrimer amino acid sequence conditions its structure, and consequently its function, would provide a considerable contribution to the overall knowledge of the field and for the development of novel applications. This poses a significant challenge since the experimental characterization of dendrimers in solution is problematic (possibly due to their high conformational flexibility [100]) and, hence, the available structural information is scarce. Computational methods, such as molecular mechanics/dynamics (MM/MD) simulations, can complement and assist in rationalizing experimental findings since they provide valuable atomic-level data. Several examples highlighting the usefulness of computational methods in the field of dendrimers are available in the literature [78, 93].

In this work, we combine experimental and computational approaches to explore the structural effects of selectively replacing positively charged residues (*L*-2,4-diaminobutyric acid, **Dab**) by neutral residues (glutamine) in the different positions of third-generation peptide dendrimers sharing a common topology and amino acid sequence (except for the mutated residues); see Figure 4.1 and Table 4.1.

Table 4.1: Peptide dendrimer sequences and synthesis.

Name	Sequence[a]	Yield mg(%)	mass[b] calc./obs.
P12	(Ac**Dab**Ser)$_8$(*Dap***Dab**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$	17(2)	3947.9/3948.9
P8	(Ac**Dab**Ser)$_8$(*Dap***Gln**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$	35(5)	4059.9/4060.9
P4	(Ac**Gln**Ser)$_8$(*Dap***Dab**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$	50(8)	4171.8/4172.8
P0[c]	(Ac**Gln**Ser)$_8$(*Dap***Gln**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$	49(6)	4284.3/4285.6

[a] Ac: acetyl group; *Dap*: *L*-2,3-diaminopropionic acid (branching residue); Amb: 4-aminomethyl-(benzoic) acid; **Dab**: *L*-2,4-diaminobutyric acid. [b] Observed by positive-ion mode MS-ESI as $[M+H]^+$, except for P0 where negative-mode was used. [c] Synthesis and experimental measurements obtained from reference 64.

We have selected Dab, a non-natural amino acid, as the positive residue to use for Dab/Gln mutations since it has the same number of side-chain aliphatic carbons as Gln and Glu. This simplifies the comparison between the results obtained here and previous ones obtained with peptide dendrimers where negatively charged residues (Glu) were also replaced by neutral ones [64, 102]. Hence, the differences between the dendrimers amino acid sequences are reduced to the side-chain functional groups (and charges) of the mutated residues: an amine group in Dab residues [–$C^{\beta}H_2$-$C^{\gamma}H_2$-NH_3^+], an amide in Gln [–$C^{\beta}H_2$-$C^{\gamma}H_2$-$CONH_2$], and a carboxylic acid in Glu [–$C^{\beta}H_2$-$C^{\gamma}H_2$-CO_2^-].

4.3 Materials and Methods

4.3.1 Dendrimer Synthesis

Solid phase peptide synthesis (SPPS) was performed manually in polypropylene syringes fitted with a polyethylene frit, a teflon stopcock and stopper. The resin (TentaGel S RAM) was swelled in dichloromethane (DCM) and the Fmoc-protecting groups of the resin were removed with a solution of 20% piperidine in DMF (2 $\times$ 10 min).

In each coupling step, Fmoc-protected amino acids (3 equiv.) and PyBOP (3 equiv.) in N-methylpyrrolidone (NMP) were added to the deprotected resin. DIPEA [N,N-Diisopropylethylamine] (5.0 equiv.) was added and the reaction was stirred for 1 hour. Reaction times were sequentially prolonged for the coupling reactions after each new branching unit: one branching unit (2h), two branching units (3h), three branch-

ing units (4h); for the branching residues coupling, twice the reaction time was used. After each coupling step the resin was washed with NMP, methanol (MeOH) and DCM (3 x each). The completion of the coupling reactions was checked using the 2,4,6-trinitrobenzensulfonic acid (TNBS) test [197]. Capping of unreacted peptide chains was ensured by using a solution of acetic anhydride/DCM (1:1, *v/v*) for 30 minutes before washing the resin again with NMP, MeOH and DCM (3 x each). Before each coupling reaction, the amino acids Fmoc-protecting groups were removed with a solution of piperidine/DMF (1:4, *v/v*) for 20 min.

At the end of the synthesis, the terminal amino groups were acetylated with a solution of acetic anhydride/DCM (1:1, *v/v*) for 1 hour. The cleavage was carried out with a TFA/H_2O/TIS/EDT (94:1:2.5:2.5) solution during 4 h. The peptide was precipitated in *tert*-butyl methyl ether (MTBE), centrifuged and dried under vacuum. Peptides were purified by preparative RP-HPLC and obtained as TFA salts after lyophilisation. The gradient used for analytical HPLC is A/D = 100/0 to 0/100 in 5 minutes, 1.2 mL min^{-1}. See Appendix C (pages 146 to 149) for further details.

4.3.2 Diffusion NMR (DOSY) Measurements and R_h calculation

Standard PSGE diffusion NMR experiments were performed using a Bruker BRX400 with dilute solutions (10 mg mL^{-1}) of dendrimer in D_2O (303 K, pH7). The pH was adjusted with dilute solutions of NaOD. The samples were prepared in air-free conditions directly before the measurements took place. The gradient with a maximum strength of 50×10^{-4} T cm^{-1} was calibrated using the HOD proton signal in D_2O (99.997%). The diffusion time Δ was 125 ms and the gradient duration δ was 5 ms. Diffusion coefficient D was derived from peak intensities using the Topspin Software from Bruker. The diffusion coefficients are average values from 14-20 analysis (intensity fit) of different ^1H signals.

The values of D are, 1.23±0.03 $m^2\,s^{-1}$, 1.27±0.02 $m^2\,s^{-1}$ and 1.32±0.02 $m^2\,s^{-1}$ for P12, P8 and P4 respectively. The diffusion coefficient and

hydrodynamic radius of P0 can be found in reference 64. The hydrodynamic radii were calculated from the diffusion coefficient D [$m^2\,s^{-1}$] using the Stokes-Einstein equation,

$$R_h = \frac{kT}{6\pi\rho D} \ , \tag{4.1}$$

where k is the Boltzmann constant ($k = 1.380 \times 10^{-23}\,J\,K^{-1}$), T is the temperature and ρ the viscosity of water at 303 K ($\rho = 1.089$ mPa s).

4.3.3 Molecular Dynamics Simulations

We have used PyMOL [172] (version 1.7.0) to obtain an initial set of 3D coordinates for each peptide dendrimer. All the dendrimers were simulated with an amino group (NH_2) attached to the core C-terminus, and acetyl (Ac) groups to the peripherical N-termini.

The GROMACS simulation package [173, 174], version 4.6.5, and the Amber ff99SB force field [198] were used to perform the MM/MD simulations. Most of the amino acid blocks considered for the topology of the peptide dendrimers were already available in the Amber ff99SB force field suite. The parameters for *L*-2,4-diaminobutyric acid (**Dab**) have been previously reported [199]. Partial atomic charges for 4-aminomethyl(benzoic) acid (Amb) and the branching *L*-2,3-diaminopropanoic acid (*Dap*) were obtained in accordance with the parameterization strategy of the force field (see Appendix C, page 150).

All simulations were done with explicit solvent, using the TIP3P water model [200] in cubic boxes, while applying periodic boundary conditions. The final systems contained approximately 54×10^3 atoms. The charges assigned to each titrable residue were the ones typically present at pH 7, namely charged Dab and charged Asp residues. The systems net charge was neutralized by adding the proper number of chlorine (eleven for P12, seven for P8 and three for P4) or sodium (one for P0) ions.

Molecular dynamics simulations were performed by integrating the equations of motion using the Verlet leapfrog algorithm with a time step of 2 fs, and the system coordinates (snapshots) were saved every 10 ps

for further analysis. The simulations were carried out in the *NPT* ensemble. The temperature was kept constant at 298.15 K using the V-rescale thermostat [201]. The pressure was kept constant at 1 atm using the Parrinello-Rahman barostat [202, 203]. A cutoff of 10 Å was used for nonbonded interactions, and long range electrostatic interactions were treated with the smooth particle mesh Ewald (PME) method [119]. Nonbonded pair lists were updated every 10 fs.

All structures were subject to energy minimization and initialization procedures previous to the production stage of the simulation. To ensure the formation of unbiased initial conformations, during some initialization steps the partial charges of all dendrimer atoms were changed from their reference values to a value of $+0.1e$ [100, 102]. This change in atomic electric charges promotes the repulsion among all dendrimer atoms, originating a generic initial structure that corresponds to the most "stretched" conformation of each dendrimer. Starting with these fully extended configurations, 10 molecular dynamics simulations (replicates) of 700 ns were performed for each dendrimer. The replicates of each peptide dendrimer were started from the same optimized system but with different sets of random velocities. The first 200 ns of each replicate simulation were considered as the systems equilibration stage and discarded. This equilibration time was determined by monitoring the radius of gyration. Only the last 500 ns of each replicate were taken as equilibrated simulations, accounting for a total concatenated production-simulation time of 5 µs (10 replicates x 0.5 µs). Only the equilibrated trajectories were used for subsequent analyses. The statistical uncertainty associated with the different quantities presented were computed using the jackknife method, leaving out one replicate on each resampling [180].

4.3.4 Asphericity

A commonly used parameter to characterize the shape of a dendrimer is its asphericity (δ) [101, 186, 188, 204–206]. This measure quantifies deviations from spherical shape and is calculated base on the three eigenvalues (λ_1, λ_2, λ_3) of the radius of gyration tensor. Asphericity is defined

as:

$$\delta = 1 - 3\frac{\langle I_2 \rangle}{\langle I_1^2 \rangle} \ , \tag{4.2}$$

where I_1 and I_2 are the first and second invariants of the radius of gyration tensor ($I_1 = \lambda_1 + \lambda_2 + \lambda_3$ and $I_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$). This quantity assumes values ranging from zero (a sphere) to one (a linear array of atoms).

4.3.5 Geometrical criteria for hydrogen bonds

The frequency of each hydrogen bond between a particular donor and acceptor was calculated simply by screening each conformation in the ensemble for the presence of the hydrogen bond and normalizing the sum by the total number of conformations. Hydrogen bonds were defined using a geometrical criterion with a maximum donor-acceptor distance of 3.5 Å and a hydrogen-donor-acceptor angle inferior to 30° [195, 196].

4.4 Results and Discussion

Dendrimers P12, P8 and P4 (Table 4.1) were synthesized by solid-phase peptide synthesis using previously described procedures [55, 64, 74]. The synthesis of P0 has been previously reported [64]. All products were obtained in good yield after purification by preparative RP-HPCL (further details in the previous section an in Appendix C, pages 147 to 149).

Standard diffusion NMR experiments were performed to determine the dendrimers diffusion coefficients (at pH7 and 303 K) and the Stokes-Einstein equation was used to obtain the corresponding hydrodynamic radius values (Table 4.2 and Figure 4.2).

The results clearly show that replacing the positively charged residues by neutral ones leads to more compact dendrimers in solution. As evidenced by Figure 4.2, this decrease in R_h values is roughly proportional to the dendrimers expected net charge, emphasizing the pivotal role played by electrostatic forces in these systems. Moreover, when

Table 4.2: Experimental and computational properties of the peptide dendrimers.

Name	Experimental		MD simulations		
	D[a] [10^{-10} m^2 s^{-1}]	R_h[b] [nm]	R_h [nm]	R_g [nm]	δ
P12	1.23	1.65±0.04	1.57±0.01	1.19±0.02	0.33±0.03
P8	1.27	1.61±0.02	1.51±0.01	1.06±0.02	0.20±0.02
P4	1.32	1.54±0.02	1.48±0.01	1.02±0.02	0.18±0.04
P0[c]	1.37	1.48±0.01	1.47±0.01	1.00±0.01	0.15±0.02

[a] Determined from dendrimer solutions in D_2O (pH 7, 303 K) using standard PSGE diffusion NMR. [b] Calculated from D using the Stokes-Einstein equation. [c] Synthesis and experimental measurements obtained from reference 64.

comparing the R_h values obtained for P12, P8 and P4 with R_h values for equivalent negatively charged dendrimers, a nearly ideal V-shaped trend is observed. This implies that dendrimers with an increasingly higher number of charged residues, whether positive or negative, adopt

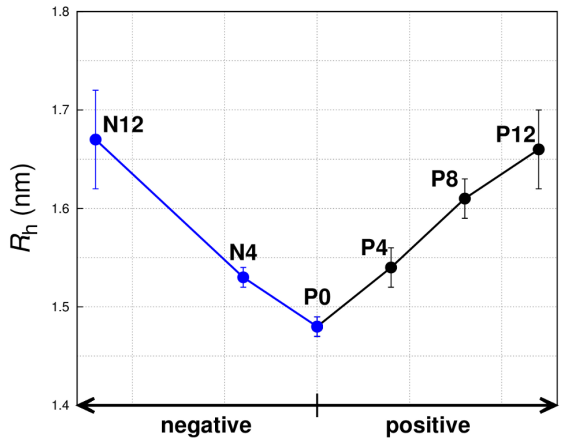


Figure 4.2: Dendrimers experimental hydrodynamic radius (R_h). Negative x-values correspond to the number of Glu residues in the sequence, positive values correspond to the number of Dab residues. The synthesis and R_h determination for dendrimers N12, N4 and P0 are described in reference 64. N12: (Ac**Glu**Ser)$_8$(*Dap***Glu**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$; N4: (Ac**Gln**-Ser)$_8$(*Dap***Glu**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$

sequentially less compact structures in solution.

To better understand the atomic-level differences between these dendrimers we have conducted long molecular dynamics simulations of P12, P8, P4 and P0. For each dendrimer we have performed a total of 5 µs production simulation using an ensemble-dynamics approach with ten replicates of 0.5 µs per dendrimer. The simulations were performed using the GROMACS 4.6.5 [173, 174] simulation package with the Amber ff99SB force field [198] in the presence of explicit solvent. The charges assigned to titrable residues were the ones typically present at pH 7 (Dab protonated, Asp deprotonated).

Computational estimates of the dendrimers hydrodynamic radius (Table 4.2) were calculated from the simulation trajectories using the HYDROPRO program [207]. The results are averages over the entire set of conformations sampled. The computed R_h values are consistent with the experimental ones, showing that the simulations adequately capture the overall compactness trend. The typical error [208] is ~9%, and most likely reflect the uncertainties associated with the estimation of hydrodynamic properties from single-solute simulations.

Another quantity frequently used to evaluate the compactness of dendrimers is their radius of gyration (R_g) [83]. In the context of dendrimers, some empirical relations between R_h and R_g have been established [101, 192, 193, 209, 210]. The calculated average R_g values are presented in Table 4.2 and, as expected, reflect the overall compactness trend for dendrimers with different amounts of charged residues.

To characterize the shape of the dendrimers we have calculated their asphericity (δ) [95, 101, 188, 204–206]. This measure quantifies deviations from spherical shape and assumes values ranging from zero (a sphere) to one (a linear array of atoms). Average asphericities are presented in Table 4.2. The results show that dendrimers bearing less charged residues are noticeably more spherically shaped. Again, a trend proportional to the number of charged residues is observed.

To analyze in detail the conformational preferences of the systems, we have resorted to two-dimensional free energy landscapes using as structural coordinates the R_g and the peptidic backbone root-mean square

deviation (rmsd) after fitting to the central structure [183] (Figure 4.3); for details on the free energy calculation see Chapter 3, section 3.3.2). Rmsd was used since its values capture the (dis)similarity among pairs of conformations. The rmsd values were computed according to a combinatorial procedure accounting for the dendrimers quasi-symmetry [100].

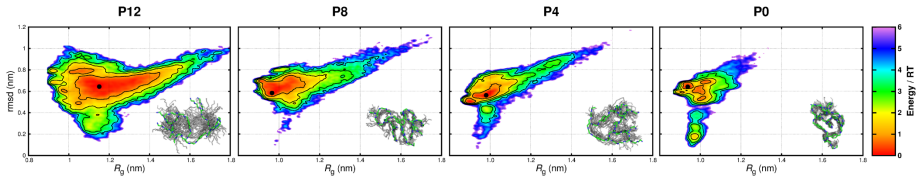


Figure 4.3: Dendrimers free energy landscapes obtained from MD simulations. The radius of gyration (R_g) and root mean-square deviation (rmsd) to the central structure are used as structural coordinates. For each dendrimer the coordinates of the lowest energy cluster are signaled by a black dot. The picture on the bottom-right margin shows the lowest energy structure (in green and blue, backbone-atoms only) superimposed on 20 other lowest energy structures from the same cluster.

The free energy landscapes in Figure 4.3 illustrate how the replacement of Dab by Gln residues in the dendrimers amino acid sequences noticeably alters the sampled conformational space. For P12 the landscape exhibits a clear downhill propensity whereas for P0, for instance, the landscape is rougher, allowing the identification of some conformational clusters. Overall, the roughness of the landscapes increases when replacing positively charged residues by neutral ones (P12→P0). A similar conformational behavior has also been previously observed in negatively charged dendrimers [102]. As expected, the lower the number of charged residues in the amino acid sequence, the lower the R_g values that are preferentially occupied. The same is observed for rmsd values, allowing us to conclude that conformational homogeneity increases in the absence of charged residues; the images of the lower energy conformers presented in Figure 4.3 reinforce this conclusion. Taken as a whole, the data presented thus far highlight the great flexibility that characterizes these systems, whereby simple modifications in key

residues side-chains can have a marked impact on the exhibited features.

Relative solvent accessibilities (RSAs) provide a useful way to identify buried and exposed residues. RSAs were computed as a residue's solvent accessibility (ASA) normalized by a reference maximum ASA value for that residue [211], and their average values are presented in Figure 4.4 (see Appendix C for details on RSA calculation, page 152). The data shows that charged residues, when present, are always the most solvent exposed (electrostatic solvation effects). Also, replacing Dab by Gln promotes the internalization of the residues in the generations where the Gln residues are present. A curious consequence is typified by the Asp residue (G0), as we go from P12 to P0 this residue becomes progressively more solvent exposed, especially in P0 where it is the sole charged residue, exemplifying the propensity to expose the available charged residues in order to achieve charge solvation.

Overall, the conformational behavior of these systems seems to result mostly from a charge balance. Although same-charge residues in a dendrimer molecule are topologically constrained to retain some proximity, the high conformational flexibility allows them to move apart in order to reduce charge–charge repulsion and increase solvation, causing the molecule to swell.

The replacement of Dab by Gln residues in the amino acid sequences modifies not only the dendrimers charge, but also the ability to establish hydrogen bonds (HBs). Dab residues present on their side-chain amine groups where N-H atom-pairs act has HB donors, whereas the side-chains of Gln residues present amide groups, and can therefore act both as HB acceptors and donors. We note that, with exception of the Dab/Gln residues, the amino acid sequences of the four dendrimers considered are equal, and hence, the HB donors and acceptors that are part of the dendrimers main-chain are identical in all the dendrimers.

In Figure 4.5 we show the average numbers of HBs split into their contributions according to the type of chain containing the donor and acceptor atoms. The data presented shows that the average number of HBs sequentially increases as Dab residues are replaced by Gln ones;

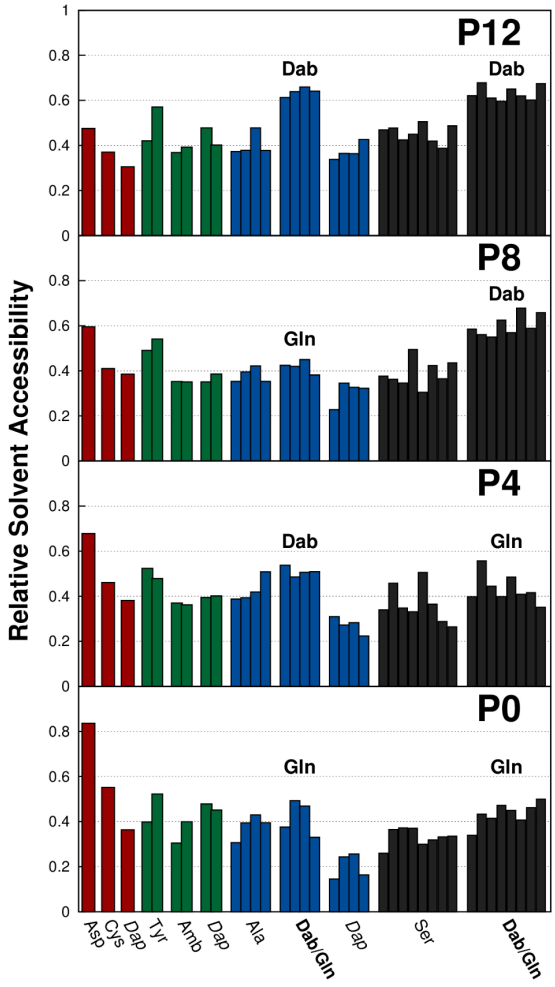


Figure 4.4: Average relative solvent accessibility of each residue obtained from the MD simulations. The residues are grouped and colored according to the generation they are part of (see Figure 4.1).

implying that an increased number of HBs is present in the more compact dendrimers and probably contributes to stabilize compact conformations. When comparing different dendrimers, the increase in the average number of HBs appears to result mostly from an increase in HBs formed between side-chain donors/acceptors and main-chain donors/acceptors and, to a lesser extent due to HBs formed between side-chains donors and acceptors. The number of HBs formed between main-chain

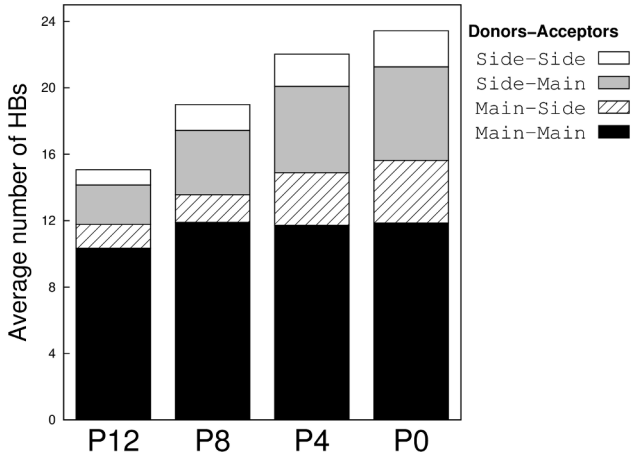


Figure 4.5: Average number of hydrogen bonds (HBs) observed in the MD simulations. Average values are partitioned according to the source of the donors and acceptors that form the bond: main-chain and/or side-chain of the dendrimer.

donors and acceptors is roughly invariant. This suggests that replacing positive residues by Gln residues mitigates repulsion effects and enables the formation of an increasingly higher number of HBs between the residue side-chain and main-chain atoms.

4.5 Conclusions

Experimental and computational methods were used in an integrated manner to provide a detailed view on the structural features of peptide dendrimers. Experimental methods were used to characterize the compactness preferences of the systems, while simulations provided a molecular-level description of the underlying phenomena. The results stress the importance of the number and placement of charged residues, and complement previous studies by showing that electrostatic forces arising from the inclusion of either positively or negatively charged residues into the amino acid sequence of peptide dendrimers promote the existence of more opened and heterogeneous structures in solution. The effects resulting from charged residues follow a trend that is pro-

portional to the dendrimers net charge.

These results can be used to great advantage to fine tune the functional behavior of dendrimers based on simple sequence design principles.

4.6 Acknowledgments

We thank Prof. Jean-Louis Reymond for helpful discussions and for hosting LCSF in his research group during the experimental part of the work. We thank also Michaela Stach, Emilyne Blattes, Gaëlle Michaud and Pedro Magalhães for helpful discussions. We acknowledge the financial support from Fundação para a Ciência e Tecnologia, Portugal, through project grants PTDC/QUI-QUI/100416/2008, UID/CBQ/-04612/2013, UID/MULTI/00612/2013, and doctoral fellowship SFRH/-BD/76085/2011. We also acknowledge the support from the COST program through action CMST/COST-Action/CM1102.

Chapter 5

Structuring peptide dendrimers through pH modulation and substrate binding

Luís C. S. Filipe[1], Sara R. R. Campos, Miguel Machuqueiro, Tamis Darbre and António M. Baptista

Contents

[1] In this work I performed all the simulations and most of the analysis.

5.1 Summary

Dendrimers are a family of ramified synthetic molecules. pH effects and electrostatic interactions are known to be crucial players to explain the conformational and functional behaviors observed in these systems. Nonetheless, to date, no computational study involving these systems has explicitly addressed the protonation equilibrium taking place at different pH values for dendrimers containing multiple ionizable sites.

Herein, we present the results of constant-pH molecular dynamics simulations performed at several pH values for four peptide dendrimers of different generations (from one to four) composed of the same type of aminoacids: histidines, serines and diaminopropionic acid. These dendrimers are known to catalyze the hydrolysis of pyrene sulfonate esters. Constant-pH MD simulations in the presence of substrate molecules at the optimum pH for catalysis are also reported.

The results show that first and second generation dendrimers are almost structurally unresponsive to pH variations. For third and fourth generation dendrimers, pH plays a structuring role, with markedly different behaviors being observed when passing from acidic to neutral pH. Protonation–conformation coupling effects influence several intramolecular interactions, which, in turn, modulate the shape and structure at the different pH values. The atypical and highly pH-dependent protonation profiles of some histidine residues are also investigated.

The interactions between dendrimers and substrates restrict the conformational space available to the dendrimers and enforces conformational homogeneity. This structuring effect is a consequence of the dendrimer–substrate interactions which occur through stabilizing hydrogen bonds and ion-pairs between the substrates sulfonate groups and the dendrimers residues.

Our results provide original fundamental data contributing to the development of novel pH-modulated dendritic systems and the improvement of the existing ones.

5.2 Introduction

Dendrimers are a family of tree-like synthetic molecules [9, 10, 16]. The ramified nature of these systems, together with the possibility of using miscellaneous building blocks in their synthesis, grants them unique structural and functional properties [7, 13, 23].

Peptide dendrimers are a particular class of dendrimers formed by alternating proteinogenic amino acids (spacer residues) with forking diaminoacids such as lysine (branching residues) [36, 37, 41, 42, 163]. The possibility of taking advantage of these dendrimers intrinsic multivalency to graft other molecules, alongside with the ability of inducing specific microenvironments by combining different amino acids into their topology, makes peptide dendrimers extremely versatile molecules [10, 163]. Such versatility has prompted the development of multiple applications, such as: metal binding dendrimers [55, 64, 65], multivalent lectin binding glycodendrimers [56, 67, 70], antimicrobial dendrimers [72, 74], cell penetrating peptides [44, 75] and, of particular relevance to the present work, catalytic dendrimers with estereolytic activity [41, 46, 47, 50, 53, 59].

A major shortcoming hindering a more extensive knowledge–based development of applications using these systems is the absence of detailed structural information at the molecular level.

Indeed, regarding dendrimers in general, the scarce crystallographic data reported in the literature has been obtained either for extremely simple systems (first generation dendrimers) [79], or recently for a dendrimer–protein complex [80] (second generation dendrimer). Therefore, it poses no surprise that computational methods have emerged as one of the main tools to explore the structural aspects of these molecules.

Despite the lack of structural information, many experimental [73, 212–214] and theoretical [83, 90, 93, 102, 185, 205, 215] studies highlight the crucial role of electrostatic interactions and pH effects on the functional and conformational behaviors of these systems.

Dendrimers are found to be quite flexible molecules [85, 93, 100, 101] able to undergo structural transitions from "dense core" to "dense shell" depending on the solution pH [83, 93, 169, 185]. For peptide den-

drimers in particular, it has recently been shown, using diffusion NMR spectroscopy, that a third generation dendrimer with multiple lysine residues becomes significantly more compact upon neutralization from pH 2 (R_h= 2.10±0.01 nm) to pH 7.4 (R_h= 1.08±0.01 nm) [74]. The direct consequence of a pH change is an alteration in the equilibrium concentrations of the protonated and deprotonated forms of the titrable sites, resulting in a corresponding change in the balance of charges, which in turn can have a direct influence on molecular structure and stability, as found in proteins [134, 135, 137, 140, 216, 217]. Therefore, the observed pH-dependent behavior of dendrimers is likely to reflect a similar tight coupling between conformation and protonation.

To the best of our knowledge, all the available computational studies exploring the pH-dependent behavior of dendrimers through molecular dynamics (MD) simulations, employ a common methodology where, depending on the specific pH one is interested in, a set of fixed protonation states are assigned to the titrable residues at the beginning of the simulation and those states are retained throughout the entire simulations. Such studies have mostly been performed for poly(amidoamine) (PAMAM) dendrimers, where the usual approach is to perform three different sets of simulations [83, 85, 169, 185, 205, 215, 218, 219]: one simulation accounting for high pH, where neither the primary nor the tertiary amines are protonated, another simulation for intermediate pH in which only the primary amines are taken as protonated and, finally, a simulation at low pH with all amines protonated. Although this approach is useful and provides valuable insights, it relies on the principle that the pK_a of the dendrimers titrable sites is similar to the "typical" pK_a of the individual sites in solution, making this approach accurate only if the titrable sites are unlikely to (de)protonate at the studied pH. Furthermore, for dendrimers containing multiple titrable sites, this approach entirely neglects the existence of site–site interactions and cooperative effects, as well as desolvation and local environment effects that influence the protonation equilibrium and can lead to large shifts in the pK_a of titrable sites.

A computational approach to address those issues is the constant-

pH MD (CpHMD) method [128, 149–154]. In CpHMD simulations, pH plays the role of an external parameter and ionizable sites can periodically protonate or deprotonate in response to the imposed pH and surrounding environment. CpHMD methods explicitly capture the tight coupling between the solute conformation and protonation states, automatically accounting for the two factors pointed above: the site–site interactions that are expected in all dendrimers rich in titrable residues; the desolvation and other local environment effects that should be significant at least in higher generation dendrimers. Therefore, CpHMD simulations seem particularly suited to study the pH-dependent conformational features of dendrimers.

Herein, we apply the stochastic titration constant-pH MD method [128, 139, 155, 156] to study four peptide dendrimers of different generations composed of the same amino acid residues (generations one to four, see Figure 5.1). Simulations were performed at six pH values, ranging from 2.5 to 7.5 with increments of one pH unit. All dendrimers are composed exclusively of histidine and serine dyads as spacer residues and *L*-2,3-diaminopropinoic acid as branching residues. Since the dendrimers terminals are capped (see Figure 5.1 and Methods section) the number of titrable residues is equal to the number of histidines in each dendrimer. Dendrimers A1, A2, A3 and A4 present 3, 7, 15 and 31 titrable residues respectively.

The dendrimers studied in this work were previously synthesized and experimentally characterized by the Reymond Group in the context of dendrimer-based catalysis [46, 47]. Those experiments showed that dendrimers A1 to A4 display esterase-type catalytic activity with histidines acting as the key catalytic residues. These dendrimers can efficiently and selectively catalyze the hydrolysis of pyrene trisulfonate esters, and a strong positive dendritic effect was observed, with both the catalytic rate and substrate binding constants increasing with the dendrimer size [46, 47]. Moreover, the experimental pH-rate profile for dendrimers A3 and A4 with 8-butyryloxypyrene-1,3,6-trisulfonate (substrate BPTS, see Figure 5.2) as substrate indicated an optimal activity between pH 5 and 6.

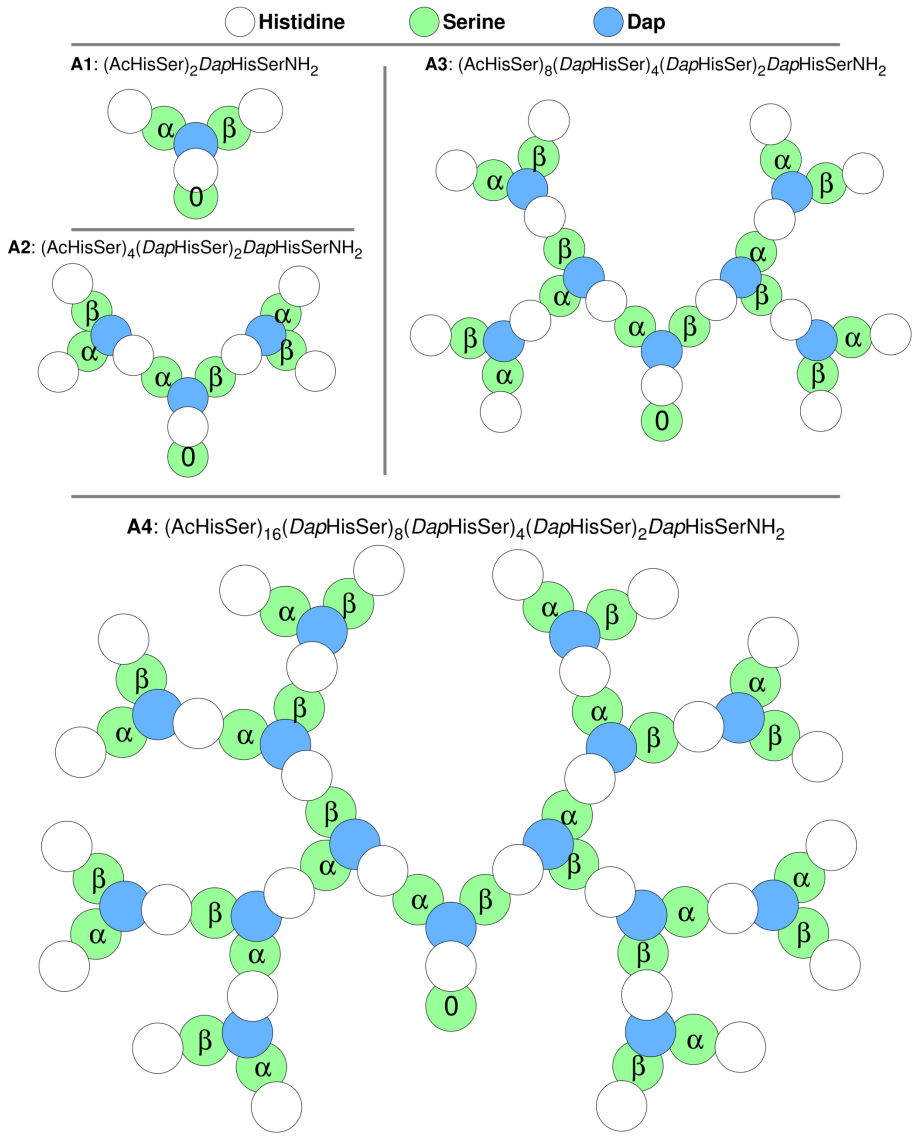


Figure 5.1: Sequence and topology of the dendrimers studied in this work (A1, A2, A3 and A4). In these dendrimers each branching diaminoacid (*Dap*: *L*-2,3-diaminopropionic acid) bifurcates the peptide sequence through two peptide bonds: one trough the N^{α}, which is the standard peptide bond observed in linear peptides; and a second peptide bond using the N^{β} from the side chain. The α and β letters in the serine circles account for the corresponding peptide bond to the preceding *Dap* residue, providing a simple way to identify individual histidine and serine residues. Each histidine/serine is named by joining the Greek letters that form the path from the dendrimer core (represented as **0**) to each specific residue. Ac = acetyl caps.

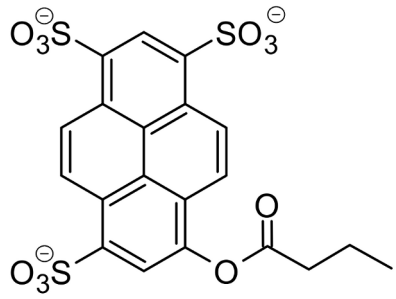


Figure 5.2: 8-butyryloxypyrene-1,3,6-trisulfonate (substrate **BPTS**).

Because electrostatic interactions can be fundamental for catalysis, and such interactions are deeply correlated with pH, we have also used the stochastic titration CpHMD method to simulate dendrimers A3 and A4 in the presence of several molecules of substrate BPTS at pH 5.5. Such simulations provide the first atomic-level description of the interactions between peptide dendrimers and substrates, unraveling the molecular basis of substrate binding in these systems.

In summary, the article is divided into two parts: in the first part we present results from simulations of dendrimers A1 to A4 in aqueous solution at different pH values, showing how the structural properties of dendrimers with multiple ionizable residues evolve with pH and with increasing generations; in the second part of the article we investigate how the presence of substrate molecules modifies the structural properties of dendrimers A3 and A4, and how the dynamics of dendrimer–substrate interactions takes place in the initial stage of the catalytic process.

To the best of our knowledge this is the first computational study of its kind, providing new information to understand the pH-dependence of peptide dendrimers.

5.3 Methods

5.3.1 Constant-pH MD Simulations at different pH values

We have performed simulations of dendrimers A1 to A4 at six pH values: 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5. For each dendrimer at each pH value, eight 100 ns long CpHMD simulations (replicates) were performed. The replicates of each dendrimer at each pH value were started from the same optimized configuration but with different sets of random velocities (see below). All production simulations were performed using the stochastic titration CpHMD method developed by Baptista and coworkers [128, 139, 155, 156]. The stochastic titration method is in essence a molecular mechanics/molecular dynamics (MM/MD) simulation where the protonation states of pre-identified titrable sites are periodically updated with new states sampled by Monte Carlo (MC) runs based on free energy terms obtained from solving the Poisson-Boltzmann (PB) equation. Algorithmically, this cyclic method relies on a stepwise approach. First, a PB/MC calculation is performed for a single solute configuration at the intended pH, with the last state of the MC simulation being taken as the new solute protonation state. The second step is a short MM/MD simulation for solvent relaxation, where the solute is kept frozen and the solvent adapts to the new solute protonation states (here, 0.2 ps). The final step is a production MM/MD simulation of the unconstrained system (here, 2 ps), whose last configuration is used as input for the following cycle.

System Setup

We have used PyMOL [172] to obtain an initial set of 3D coordinates for each dendrimer depicted in Figure 5.1. All the dendrimers were simulated with an amine group attached to the core C-terminus and acetyl groups to the peripheral N-termini.

MM/MD Settings

The GROMACS package [173, 174], version 4.0.7, the GROMOS 54A7 force field [115] and single point charge (SPC) water molecules [178] were used to perform the MM/MD simulations. The topology parameters for serine and the different protonation forms of histidine are already included in the GROMOS 54A7 standard set [115]. The topology block for *L*-2,3-diaminopropanoic acid (Dap) has previously been reported for the GROMOS 53A6 force field [100], and was adapted by modifying the peptide bond dihedral angles (in line with the GROMOS 54A7 definition).

All the initial structures were subjected to energy minimization and initiation procedures using standard MD simulations previous to the production stage with CpHMD simulation. The initiation protocol is similar to the one used in references 100 and 102. Briefly, the initial configurations were first energy minimized and then subjected to several standard MD initiation runs in vacuum. To ensure the formation of unbiased initial conformations, during some initiation steps the partial charges of all the dendrimers atoms were changed from their reference value to a value of +0.1 e. This change in atomic charges promotes the repulsion among all dendrimer branches, originating a generic initial structure that corresponds to the most "stretched" conformation of each dendrimer. The systems were then solvated, energy minimized again to remove excessive strain, and a new set of standard MD initiation runs were performed to allow the systems to relax and adapt. The aforementioned initiation runs were performed eight times using different random velocities, originating the eight replicates performed for each dendrimer at each pH value.

All production simulations were done with explicit solvent in rhombic dodecahedral boxes, while applying periodic boundary conditions. The final systems contained approximately 15×10^3, 27.5×10^3, 40×10^3 and 94×10^3 atoms for A1, A2, A3 and A4 respectively. The equations of motion were solved using the leapfrog algorithm with a time step of 2 fs, and the systems coordinates (snapshots) were saved every 10 ps. The LINCS algorithm [220] was employed to keep all bonds at their

equilibrium values and the SETTLE algorithm [221] was used to maintain water molecules rigid. Temperature and pressure were kept constant at 298.15 K and 1 bar by coupling to external baths [177]. Nonbonded interactions were treated with a twin-range cutoff of 8/14 Å and the neighbor lists updated every 10 fs. The generalized reaction-field method [117] with a relative dielectric constant of 54 [176] and ionic strength of 3 mM (A1), 7 mM (A2), 15 mM (A3) and 31 mM (A4), estimated from the experimental conditions [47], was used for the long-range electrostatic interactions.

PB/MC Settings

PB and MC calculations were performed with the MEAD [194] (version 2.2.9) and PETIT [148] (version 1.5) packages, respectively. Protonation states were saved every 2 ps. The atomic charges and radii were taken from the GROMOS 54A7 force field as described elsewhere [143]. The model compound pK_a used (6.88 for histidine) and its calibration method are reported elsewhere [222].

All PB calculations consisted of finite-difference linear PB calculations using a temperature of 298.15 K, a molecular surface defined with a solvent probe radius of 1.4 Å, and a Stern (ion exclusion) layer of 2.0 Å. The dielectric constants were 80 for solvent and 2 for the dendrimers. A two-step focusing procedure [223] was used, with consecutive grid spacings of 1.0 Å and 0.25 Å.

The MC runs were performed using 10^5 MC cycles, one cycle consisting of sequential state changes (including tautomeric forms) over all individual sites and also all pairs of sites with at least one interaction term above 2.0 pK_a units [133, 148], whose acceptance/rejection followed a Metropolis criterion [125]. All histidine residues were treated as titrable sites.

Equilibration times

Each replicate equilibration time was chosen by analyzing the dendrimer radius of gyration (R_g) and total charge time-series. R_g provides a mea-

sure of dendrimer size and is widely used in computational studies of dendrimers [83, 101, 102, 185, 186, 188]. The cumulative average protonation of individual histidine residues at different pH values was also taken into account. The replicates sampling convergence and convergence amongst replicates was also evaluated (Appendix D, pages 154 to 157). The systems showed to be equilibrated at different time lengths and only the equilibrated trajectories were used for subsequent analysis. The length of the composite trajectories obtained by concatenating the equilibrated portions of each set of 8 replicates ranged between 575 to 645 ns for A1, 525 to 580 ns for A2, 470 to 530 ns for A3 and 410 to 480 ns for A4.

5.3.2 Constant-pH MD Simulations with substrates

We have performed CpHMD simulations of dendrimers A3 and A4 at pH5.5 (the optimal pH for catalysis [46, 47]) in the presence of 5 (for A3) and 10 (for A4) molecules of 8-butyryloxypyrene-1,3,6-trisulfonate (substrate BPTS, see Figure 5.2). The number of substrate molecules was chosen based on results from isothermal titration calorimetry experiments showing that, depending on the pyrene sulfonate compound used in the experiment (Table 2 in reference 46 and Table 1 in reference 47), A3 has 4.78 to 5.83 binding sites, and A4 has 9.43 to 9.88 binding sites. The force field parameters used for substrate BPTS are presented in Appendix D (pages 158 to 163).

For each dendrimer, initial configurations were randomly selected from the equilibrated trajectories of the simulations at pH 5.5 (see the previous subsection) and substrate molecules were added randomly to the simulation box, while ensuring that all molecules were placed far from the dendrimer. After solvation and energy minimization procedures, five initiation simulations with different random initial velocities (replicates) were performed using standard MD, followed by 120 ns long CpHMD simulations. The final systems contained approximately 51×10^3 and 95×10^3 atoms for A3 and A4 respectively.

The system coordinates were saved at 1 ps intervals and the protonation states resulting from the PB/MC steps were saved every 2 ps.

Since sulfonic groups are salts of strong acids and can be considered as negatively charged in any relevant pH conditions (pK_a for sulfonic acids is < 0), only the dendrimers histidine residues were considered as titrable sites. Substrate molecules were included in the PB/MC calculations. The other MM/MD and PB/MC settings follow what has been described for the CpHMD simulations without substrate (previous subsection).

Each replicate equilibration time was chosen based on the analysis of three properties: R_g, total charge and the sum of all the substrate molecules contact area with the dendrimers. The corresponding plots and sampling convergence analysis are presented in Appendix D (pages 165 to 167). The final concatenated trajectories were 410 ns long for A3 and 360 ns for A4.

5.3.3 Analyses

Asphericity

See Chapter 4, section 4.3.4.

Geometric criteria

Hydrogen bonds (HBs) were defined using a geometric criterion with a maximum donor–acceptor distance of 3.5 Å and a hydrogen–donor–acceptor angle inferior to 30° [195, 196]. In the context of peptidic systems containing histidines in different protonation states, hydrogen bonds are particularly relevant, since the imidazole nitrogen atoms are both HB donors in the charged form, whereas one of them becomes an HB acceptor in the neutral form.

The side chain imidazole of histidine is an aromatic motif; hence histidine residues can form stacking interactions [224–226] even when both histidines are charged [227]. Stacking between two histidines imidazole rings was determined based on the distance between the centroids of both rings (if < 5.2 Å), the horizontal displacement between both centroids after orientation in a common reference geometric frame (if < 4.5 Å), and the angle formed between the surface normal vectors of

the planes defined by each imidazole ring (Γ, the normal–normal angle) [225, 227, 228]. We employ a conservative set of criteria that considers a stacking interaction only if the two ring planes are either parallel ($150° < \Gamma < 30°$, encompassing parallel stacked and parallel displaced geometries) or perpendicular ($60° < \Gamma < 120°$; encompassing edge-to-face geometries) [224].

Hydrogen bonding is a broad phenomenon, and a particular case not accounted in our initial criteria (see above) is X-H···π bonds. X-H···π interactions occur between a non-aromatic X-H moiety and an aromatic ring (acceptor) [229, 230, 230]. X-H···π interactions were determined using the following criteria: the distance between the imidazole centroid and the X donor-atom (donor···COM < 4.0 Å); the distance between the imidazole centroid and the H atom (H···COM < 3.0 Å); and the X-H-centroid angle (X-H···COM > 120°) [225, 226, 229]. Only N-H···π and O-H···π interactions were considered.

Experimentally it is known that X-H···π interaction can occur only if a histidine side-chain is in its neutral form [226, 229]. We tested the robustness of the criteria employed by computing X-H···π interactions considering all histidine residues in the calculations, regardless of their charged form, and performing the same calculations while considering only neutral histidines. The results show that the number of X-H···π bonds found is identical in both cases (see Appendix D, page 167); reinforcing the suitability of the criteria used and the ability of our simulations to reproduce experimental data.

The formation of ion-pairs between charged histidine residues and the sulfonate groups of substrate BPTS was considered to occur when the distance between the charged groups was smaller than 3.5 Å [231, 232]. If a histidine residue is neutral then its imidazole ring can still interact with the sulfonate groups of BPTS through hydrogen bonds (with the substrate oxygen atoms acting as acceptors and the imidazole protonated nitrogen as donor).

Energy landscapes

See Chapter 3, section 3.3.2.

Errors

The statistical uncertainties associated with the different quantities presented were computed using the jackknife method [180], which is a leave-one-out resampling strategy. In particular, instead of computing the value of a property using all the 8 replicates available, one replicate was excluded from the dataset in turn and an average value is computed using the remaining 7 replicates. The statistical uncertainty is then calculated as the jackknife standard error between the value obtained using the entire dataset (all the replicates) and the 8 values obtained when excluding one replicate at a turn.

The sole exception were the errors of the histidines pK_a values, which were computed using a bootstrap approach [180] identical to the one presented in reference 233, but interpolating the data with a cubic spline [234] instead of fitting it to the Hill equation as done in the aforementioned reference.

5.4 Results and Discussion

5.4.1 Peptide dendrimers at different pH values

Charge, size and shape

The dendrimers protonation curves obtained from the simulations are in excellent agreement with the experimental data (Figure 5.3, top), ensuring that the simulations adequately reflect the experimental systems.

Regarding size and compactness, Figure 5.3 (middle plot) illustrates how dendrimers of different generations but composed of the same kind of amino acids can have markedly different behaviors. The higher the generation, the more dependent on pH its R_g becomes.

Although, the size of the smallest dendrimer, A1, remains mostly unaffected by pH, a common trend is observed for A2, A3 and A4, namely that they have their higher R_g values at the lowest pH. This is most likely a consequence of the preponderance of electrostatic repulsion effects between charged residues, since that is the pH at which more histidines are in their charged form; obviously such repulsion ef-

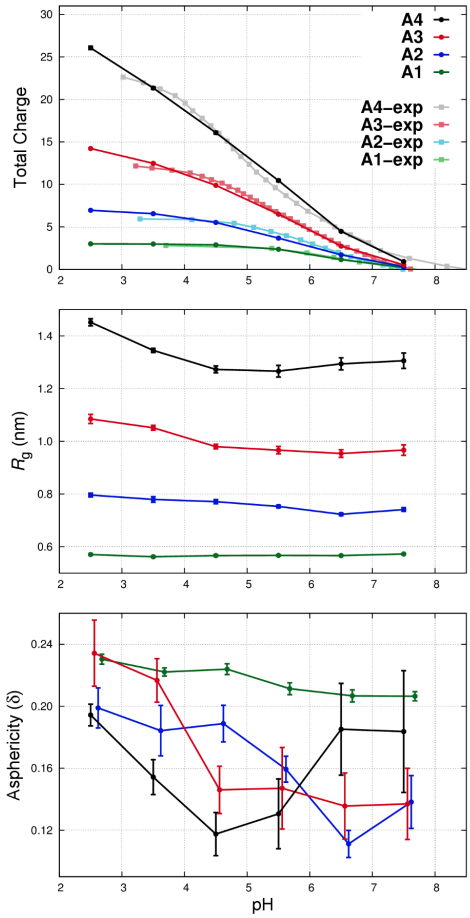


Figure 5.3: (*Top*) Experimental and computational protonation curves. The experimental data for A3 and A4 was retrieved from reference 47; the data for A1 and A2 is presented as supporting information (see Appendix D, pages 168 to 170). (*Middle*) Average radius of gyration at different pH values. (*Bottom*) Asphericity at different pH values; pH values are slightly shifted for better visualization of the error bars.

fects are attenuated with increasing pH. It is noteworthy that other authors who performed simulations of different dendrimers considering distinct overall charged states, also observe a consistent trend where the dendrimers size is higher at lower pH values [83, 185]. However, while A2 and A3 show a systematic size decrease with pH (more sigmoidal in the latter), A4 acquires its lowest compactness around pH 5. Since

A4 is considerably larger than the other dendrimers, this may reflect the emergence of subtler conformation-protonation couplings besides electrostatic repulsion (see next section).

The bottom plot of Figure 5.3 allows us to infer the pH dependence of dendrimer shape. It can be seen that A1, A2 and A3 tend to become more spherical near neutral pH. However, for A4 this shape descriptor trend is similar to the one observed for its R_g values, with the lowest asphericity values observed at pH 4.5 and 5.5. Again, δ reflects a remarkable influence of pH on these systems shape acquisition. When compared with, for instance, δ values observed for different generation PAMAM dendrimers [188] (which range from 0.15 to 0.03) the peptide dendrimers studied here are considerably more aspherical. This is not surprising considering the intrinsic asymmetry of peptide dendrimers and the fully symmetric nature of PAMAM dendrimers. Nevertheless, the values of δ observed here are very similar to the ones estimated by Falkovich and coworkers from simulations of peptide dendrimers ranging from generation one to five and composed solely of lysine residues (with δ between 0.35 and 0.1) [101].

Conformational preferences

As the previous results indicate, pH has a considerable effect on the preferential conformations adopted by peptide dendrimers, an effect that is especially noticeable in higher generation dendrimers. To analyze the conformational preferences of these systems we resorted to two-dimensional energy landscapes using R_g and SASA as structural coordinates. Probability density histograms of R_g and average values of SASA are presented in Appendix D (pages 171 and 172).

To illustrate the behaviors observed, we present in Figure 5.4 the energy landscapes of A4. As shown, the roughness of the landscapes increases with increasing pH, and distinct clusters of structurally similar conformations can be identified at pH 5.5, 6.5 and 7.5. It is significant that no clear conformational clusters can be identified at low pH values, with the lowest energy structures being quite heterogeneous. Thus, our results indicate that increasing pH promotes the partitioning into con-

formationally more homogeneous clusters. Furthermore, from pH 4.5 to 7.5 both the R_g and SASA values seem to be restricted to a smaller (but still pH-dependent) range. Overall, this can be described as a pH-induced structuring effect.

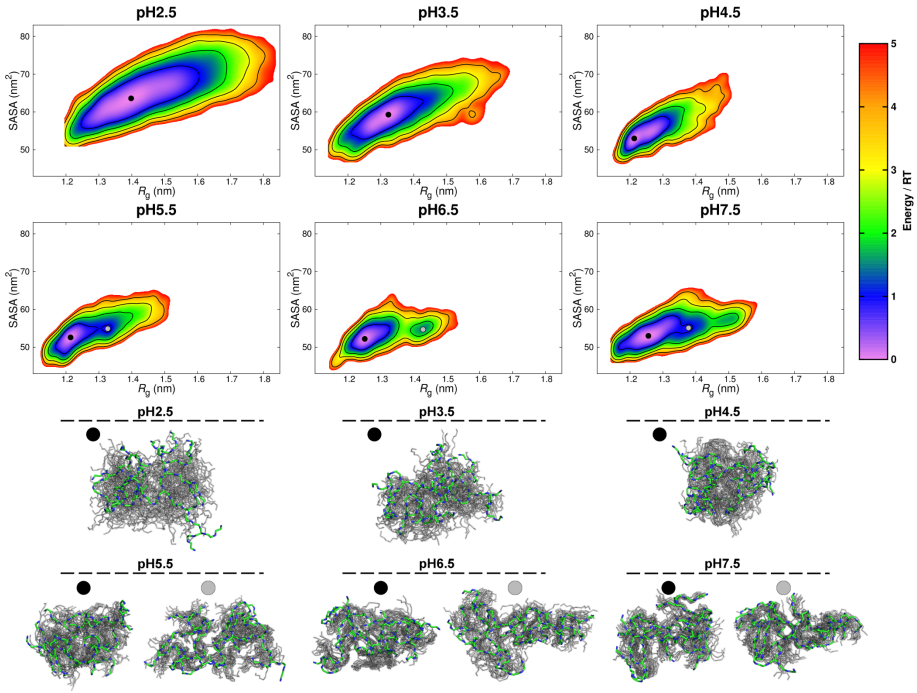


Figure 5.4: Dendrimer A4 energy profiles and lowest energy conformers at different pH values. The energy profiles use R_g and SASA as structural coordinates. In the energy profiles the lowest energy cluster is represented as a black dot. Pictures of the 21 lowest energy conformers of each cluster are shown, with the lowest energy structure highlighted (only the peptidic backbone is represented using green for carbons and blue for nitrogen atoms) and superimposed on the remaining 20 structures (in gray).

When structural clusters are identifiable (above pH 4.5), the lowest energy cluster (black circle) always accounts for very compact structures with all the different dendrimer branches in close proximity, whereas local conformational clusters (gray circles) seem to account for more open conformations with specific interaction patterns between the different parts of the dendrimer.

The energy landscapes of A1, A2 and A3 are presented in Appendix D (pages 173 to 175). The pH-dependent conformational behavior of A3 is essentially similar to the one of A4. However, the landscapes of A1 and A2 show only a downhill propensity with no identifiable conformational clusters. These results reflect the striking difference in the number of amino acids that compose each dendrimer (7, 17, 37 and 77 for A1, A2, A3 and A4 respectively), with only higher generation dendrimers being complex enough to exhibit structured regions, while lower generation dendrimers are very flexible and highly exposed to the solvent. A similar behavior was observed by Chalmers and coworkers [95] in simulations of poly(L-lysine) dendrimers of generations one to six; the authors described this type of transition as being a change from a small molecule-like behavior, for lower generations dendrimers, to a polymer-like behavior in higher generation dendrimers.

Specific interactions at different pH values

The high content of histidine residues, whose imidazole groups can be charged or neutral, makes these peptide dendrimers extremely versatile in terms of the interactions they can establish: charged imidazole groups can form ion–pairs with molecules of opposite charge (e.g. substrate BPTS); the imidazole protonated nitrogens can act as hydrogen-bond donors whereas the basic nitrogen atom (present in neutral histidines) can act as hydrogen-bond acceptor; the aromatic character of the imidazole ring allows the formation of stacking interactions with other aromatic rings (including imidazoles from other histidines); and hydrogen-π interactions (X-H$\cdots\pi$) between the imidazole ring and polar hydrogens of non-aromatic donors, such as serine side-chain hydroxyl group or main-chain nitrogen atoms, can also exist. To define the existence of each of these interactions we applied specific geometrical criteria (see the Methods section). Some of the interactions are mutually exclusive; for example, if two histidines are considered stacked, hydrogen bonds cannot be formed between their imidazole atoms.

As Figure 5.5 (top) illustrates, the higher the dendrimer generation the more preponderant π-π interactions become. Also, these interac-

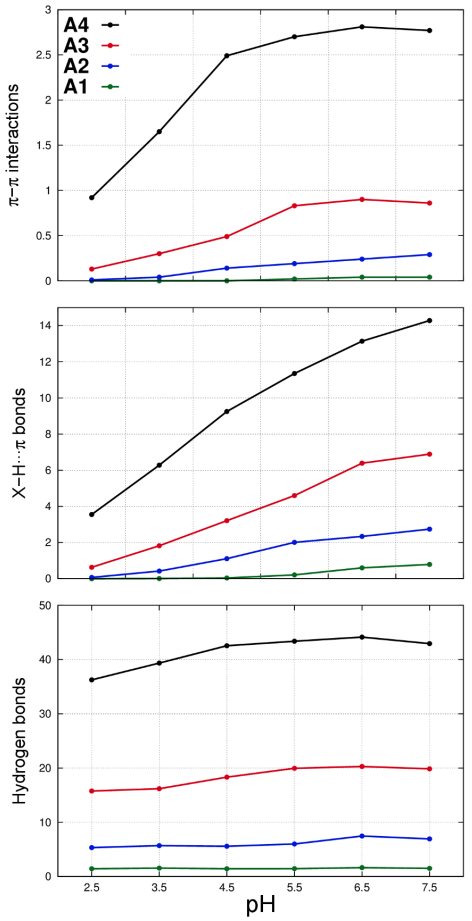


Figure 5.5: pH-dependent interactions of dendrimers A1 to A4. (*Top*) Average number of π-π interactions. (*Middle*) Average number of X-H···π bonds. (*Bottom*) Average number of hydrogen bonds.

tions seem to occur preferentially at pH > 5. Indeed, for dendrimers A3 and A4, where such interactions are more frequent, the average number of π-π interactions remains roughly stable between pH 5.5 and pH 7.5. A detailed histogram analysis (Appendix D, page 176) reveals that such interactions rarely occur between charged histidine residues, taking place almost exclusively between a charged and a neutral histidine or between two neutral histidines; with the later becoming more prevalent as pH increases. Moreover, the majority of π-π interactions found in these dendrimers are interactions where the histidines imidazole rings

are perpendicular to one another (edge-to-face stacking), whereas the "typical" parallel (or parallel displaced) stacking corresponds only to 20% of the π-π interactions observed.

Figure 5.5 also shows the average number of X-H···π bonds (middle plot). X-H···π interactions occur between a non-aromatic X-H moiety and an aromatic ring (acceptor). In this type of interaction the donor atom X is placed roughly above the center of an aromatic ring, with the X-H bond pointing at it [229, 230]. O-H···π and N-H···π interactions are known to regularly occur in proteins [230]. These interactions are present in all dendrimers studied in this work, and their average number increases with the dendrimer generation. Also, as pH increases these interactions become more frequent, in an almost linear trend that is shared by the four dendrimers. Such trend is easily explained by the fact that, as seen in Figure 5.3, the number of charged histidines diminishes in an almost linear fashion and X-H···π bonds can only occur if a histidine side-chain is in its neutral form (because cationic imidazole is unsuitable as a hydrogen bond acceptor [226, 229]).

Regarding hydrogen bonds (Figure 5.5, bottom) the trend observed is similar to the one described for π-π interactions. Our simulations show that for A3 and A4 the number of hydrogen bonds increases as pH changes from 2.5 to 5.5, with the average number of HBs remaining stable between pH5.5 and pH7.5. We have analyzed which parts of the dendrimers are preferentially involved in the formation of HBs (see Appendix D, page 177) concluding that the biggest contribution to the total number of HBs comes, in all dendrimers, from bonds formed between main chain acceptor and donor atoms. Hydrogen bonds where one of the participants belongs to the main chain and the other to a side chain are also regularly observed, whilst HBs where both participants belong to side chains are less frequent. We had previously observed this kind of HB pattern in simulations of third generation peptide dendrimers containing variable amounts of charged residues [102]. It was our rationale then, as it is now, that a balance between electrostatic interactions promoting repulsion among charged residues, and the network of main chain HB that becomes preponderant when repulsion effects are attenu-

ated, are key structural determinants in these systems. Understandably, in systems such as the ones studied here, where multiple histidines or other aromatic residues are present, X-H···π and π-π interactions need also to be taken into account.

Individual protonation curves and pK_a values

As previously discussed, dendrimers A1 and A2 are quite flexible and no clear preferential set of conformations is identifiable at any particular pH value. The pK_a values computed for the histidines of these dendrimers reflect that feature (see Figure 5.6), with the pK_a values distributed within a small range (6.06–6.59 for A1 and 5.10–5.93 for A2) regardless of the dendrimer shell they belong to. Moreover, although those values are lower than the pK_a of 6.5 found in solution [235], as one would expect from charge repulsion, the shifts are small, evidencing the fact that histidine residues are mostly solvent exposed.

As we look into the more complex dendrimers A3 and A4, where preferential conformations occur at certain pH values, different individual protonation behaviors emerge. Broadly, we can classify the protonation curves computed for A3 and A4 as either sigmoidal or non-sigmoidal.

Sigmoidal curves are observed for most histidines of A3 and A4, as exemplified by His$^{\beta\alpha\beta}$ of A3 in Figure 5.7. Non-sigmoidal curves, with a localized small peak or plateau, are observed for a few histidines of A3 and A4, as exemplified by His$^{\beta}$ of A3 in Figure 5.7. This localized small peak or plateau is usually associated with an increased solvent exposure of the imidazole ring, as also illustrated in Figure 5.7: while the ring of His$^{\beta}$ of A3 is substantially buried at pH 5.5, it becomes more exposed at pH 6.5 and 7.5, which provides solvation for the charged form and shields it from the repulsion by other charged histidines, thereby allowing an increased protonation despite the higher pH. These non-sigmoidal protonation curves also show how a single pK_a value can be a rather poor way to characterize a titrable group in a complex system, especially when pH-induced conformational effects take place.

The full set of protonation curves and imidazole SASA are presented

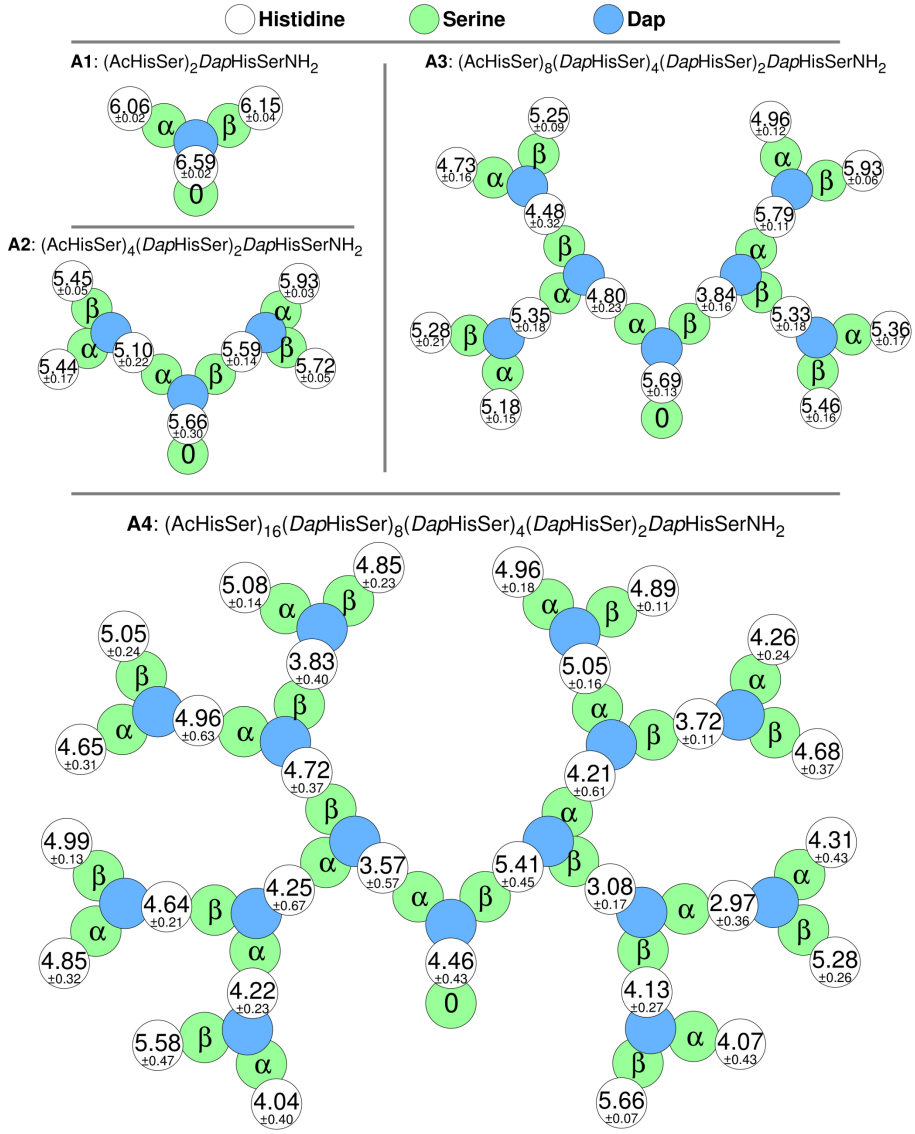


Figure 5.6: Histidine residues pK_a values and associated errors. For further details see Figure 5.1.

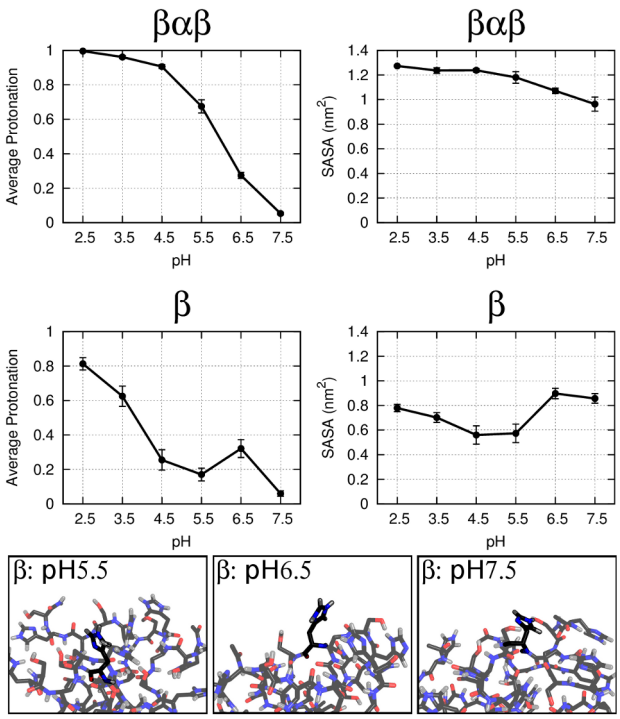


Figure 5.7: (*Top*) Average proton occupancy and imidazole ring solvent accessible surface area curves for $His^{\beta\alpha\beta}$ (pK_a= 5.93±0.06 in Figure 5.6) and His^{β} (pK_a= 3.84±0.16 in Figure 5.6) of A3. (*Bottom*) Pictures from the lowest energy structures of A3 energy landscape (provided in Appendix D, page 175) highlighting His^{β} and its surroundings at pH5.5, 6.5 and 7.5.

as Appendix D (pages 178 to 185).

5.4.2 A3 and A4 simulations with substrates

Substrate–induced structural changes

We have performed CpHMD simulations of dendrimers A3 and A4 at pH5.5 (the optimal pH for catalysis) in the presence of 5 (for A3) and 10 (for A4) molecules of substrate BPTS (Figure 5.2). These numbers of substrate molecules were chosen based on results from isothermal titration calorimetry (ITC) experiments (see Methods) [46, 47].

A substrate molecule was considered to interact with the dendrimer

Table 5.1: Average dendrimer charge, radius of gyration and asphericity at pH 5.5 in the absence and presence of substrate molecules.

	A3		A4	
	no substrates	5 substrates	no substrates	10 substrates
Total charge	6.49 ±0.06	11.66 ±0.28	10.45 ±0.12	23.33 ±0.57
R_g (nm)	0.97 ±0.01	0.92 ±0.01	1.27 ±0.02	1.24 ±0.02
δ	0.15 ±0.03	0.05 ±0.01	0.13 ±0.02	0.05 ±0.02

if at least 15% of its total surface area is in contact with the surface of the dendrimer. In Appendix D (pages 185 and 186) we detail the specifics of how the contact area was computed, alongside with histograms of contact area distributions and tests showing the robustness of the cutoff chosen.

Experimentally it has been shown that, depending on the pyrene sulfonate compound used, dendrimer A3 has 4.78 to 5.83 binding sites, and A4 has 9.43 to 9.88 binding sites (ITC experiments using BPTS were not performed) [46, 47]. Here, we observe that, on average, 4.97±0.14 and 9.68±0.15 molecules of substrate interact simultaneously with dendrimers A3 and A4 respectively. Dendrimers A3 and A4 can indeed simultaneously accommodate 5 and 10 substrate molecules. Representative snapshots of dendrimer–substrate complexes are presented in Appendix D (page 187). Due to the difficulty of capturing the substrate release dynamics in the simulated time scale, our simulations show that substrate molecules are normally interacting with the dendrimer; when a substrate–dendrimer contact area decreases below the cutoff, it reflects the transient interaction of that substrate molecule with another one in contact with the dendrimer. This kind of behavior has previously been observed by others in simulations of dendrimer–guest systems with multiple guests [90, 236, 237].

As shown in Table 5.1, the presence of substrate molecules considerably modifies the average charge and shape of peptide dendrimers. Regarding the dendrimers total charge, our results show that on average approximately 75% of the histidine residues are charged when substrates are present (as opposed to 43% and 34% in the absence of sub-

strate at pH 5.5 for dendrimers A3 and A4 respectively). These results are consistent with the available experimental data, namely titration curves of A3 and A4 in the presence of 5 and 10 equivalents of 1,3,6,8-pyrene tetrasulfonate. In those experiments, and comparing with titration curves in the absence of substrates, a considerable shift in the curves towards higher pH values was observed for both A3 and A4 [46, 47]. Such increased protonation is what one would expect from the electrostatic effect of the bound negative substrates.

Figure 5.8 highlights the changes in the histidines proton affinities that take place when substrate molecules are present. Besides the marked difference in the total number of histidines that preferentially adopt a charged form, one can see that most neutral histidines were also found frequently in the neutral form in the simulations without substrates. In general, histidines that are prone to be in their neutral form are consistently less solvent exposed, a clear consequence of the structure adopted by these dendrimers, which probably keep those histidines in a peptidic environment where they are inaccessible to the solvent and remain deprotonated by the influence of neighbor charged histidines.

As seen from the R_g values in Table 5.1, when substrate molecules are present, peptide dendrimers become slightly more compact, and the histograms become more peaked (compare Figures D.13 in page 171 and D.8, page 167). Their average asphericity shows that the conformations acquired become noticeably more spherical when compared to the ones sampled in the simulations without substrates.

Comparing the energy landscapes of A3 and A4 in the presence (Figure 5.9) and absence (Figure 5.4 and Figure D.17 in Appendix D, page 175) of substrate molecules further supports this increased compactness. In the presence of substrates the sampled conformations occupy a smaller area of the plot, implying that substrate molecules considerably restrict the conformational space accessible to these dendrimers, rendering them less flexible. Furthermore, Figure 5.9 shows that substrates enforce the existence of clearly defined energy wells accounting for clusters of very homogeneous conformations. Thus, substrate binding has an overall structuring effect on these peptide dendrimers.

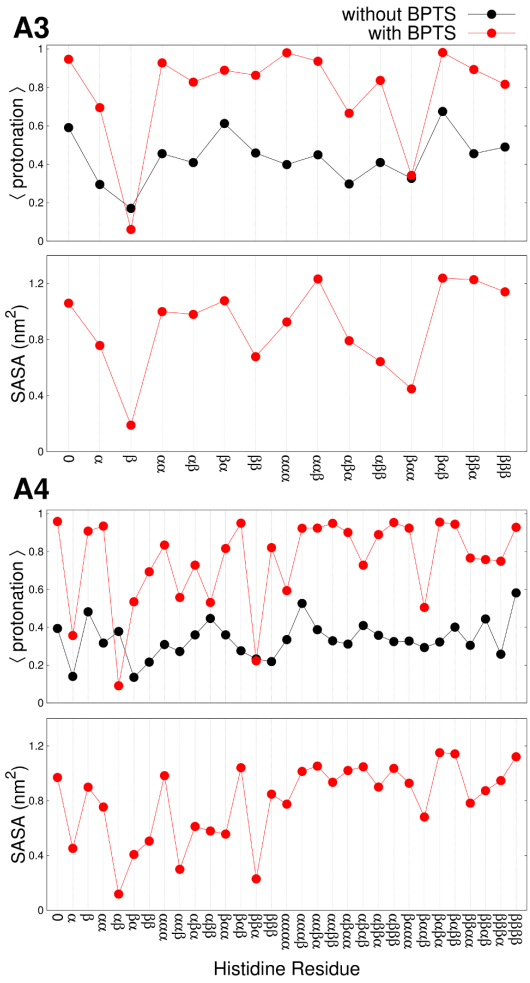


Figure 5.8: Histidine average protonation computed from simulations in the absence and presence of substrate molecules at pH 5.5 and imidazole solvent accessible surface area (SASA) computed from the simulations with substrates but considering only the dendrimer. Dendrimer A3 (*top*) and A4 (*bottom*). See Figure 5.1 for details on histidine residue naming.

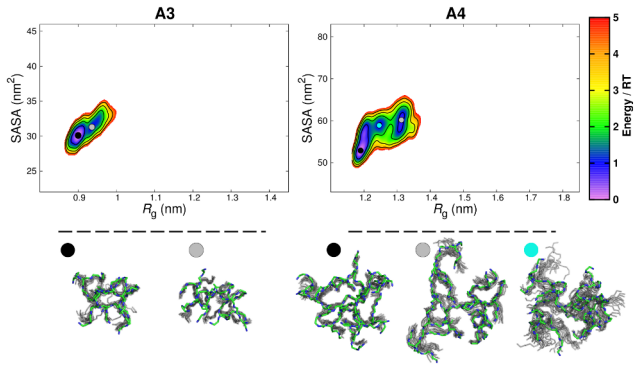


Figure 5.9: A3 and A4 energy profiles in the presence of 5 and 10 molecules of BPTS. The lowest energy cluster is represented by a black dot. The lowest energy structure of each cluster is represented colored in green and blue, and twenty other low energy conformations are drawn in gray (peptidic backbone only). The two plots are drawn using different scales to make the comparison with the simulations without substrates easier (Figure 5.4 and Figure D.17 in Appendix D, page 175).

Dendrimer–Substrate interactions

Radial distribution functions of dendrimer–substrate pairs show that substrate molecules are mostly found at distances above 9 Å of the dendrimers center-of-mass (see Figure D.27 in Appendix D, 187). However, as Figure 5.10 (left) emphasizes, several residues are close to substrate molecules. The emerging picture is that substrate molecules adhere to the dendrimer surface and partially penetrate it, without getting deeply buried (further supported by dendrimer–substrate contact area histograms, Figure D.24 in Appendix D, page 186). Additionally, we have found no evidence for the incorporation of the substrate aliphatic tails into the dendrimer interior, contrary to what might be expected from their hydrophobic nature (see histograms of the dendrimer-aliphatic chains contact areas in Figure D.27 of Appendix D, page 187).

Indeed, our results suggest that the initial steps of the catalytic process take place close to the surface of the dendrimer, where the residues that are near substrate molecules form a network of transient hydrogen bonds and ion-pairs that stabilize the interactions (Figure 5.10, middle and right). If on one hand, ion-pairs can only be formed between the

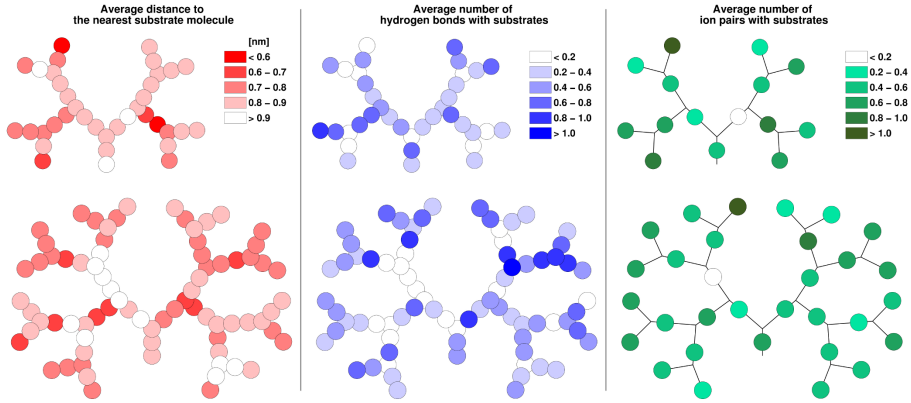


Figure 5.10: Topological representation of the interactions between A3 and A4 with substrate molecules. (*Left*) Average distance between the residue center-of-mass (COM) and the COM of the nearest substrate molecule. (*Middle*) Average number of hydrogen bonds formed between the dendrimer and the substrates. (*Right*) Average number of ion-pairs formed between the dendrimer and the substrate; only histidine residues are represented since only charged histidines from the dendrimer can establish ion-pairs with the sulfonate groups of the substrates.

dendrimer charged histidine residues and the sulfonate groups of the substrates, hydrogen bonds can be formed between any dendrimer HB donors and the HB acceptors of the substrate sulfonate and ester groups. The data presented in Table 5.2 shows that sulfonate acceptors from the substrate are the main responsible for the formation of HB with the dendrimers; which seems a reasonable result since the ester group should stay available in order to be hydrolyzed during catalysis. The fact that, in both dendrimers, substrates establish roughly the same number of HB with the dendrimer main chain and side chain donors, suggests that the formation of dendrimer–substrate HB is unspecific; possibly the substrate just requires some anchoring HB donors in the dendrimer.

Table 5.2: Average number of hydrogen bonds formed between different parts of the dendrimer (donor atoms) and substrates (acceptor atoms).

	A3		A4	
Substrate	main chain	side chain	main chain	side chain
Sulfonate groups	6.44	6.70	16.55	14.66
Ester groups	0.43	0.20	0.97	0.36
Substrates (total)	13.78		32.54	

5.5 Conclusions

Constant-pH MD simulations were used to provide a complete characterization of four dendrimers of different generations. By using a computational method that explicitly captures the coupling between protonation and conformation, we were able to study in detail the behavior of the dendrimers at different pH values, and disclose features of these systems that were previously unknown. Furthermore, the simulation results are consistent with the available experimental data.

We observe that for lower generation dendrimers (A1 and A2) only minor conformational changes take place when passing from acidic to neutral pH. The protonation curves of histidine residues in those dendrimers reflect the dendrimers high flexibility and solvent exposure of the imidazole rings. For higher generation dendrimers (A3 and A4), the data presented unravels a more complex behavior where distinct conformational preferences arise above pH 4.5, leading to more structured molecules. In general, these dendrimers histidine residues protonate at lower pH values than the ones observed for A1 and A2, reflecting the complex relations established by the different titrable residues, including the existence of desolvated histidine residues buried in a peptidic environment. However, at specific pH values, conformational changes that expose the aforementioned residues to the solvent are possible.

We have also explored the interactions between higher generation dendrimers (A3 and A4) and the substrate molecules involved in dendrimer–mediated catalysis. Our results provide the first atomic-

level glimpse into how the presence of substrates affects the protonation and conformational behavior of peptide dendrimers at the optimum pH for catalysis, namely by restricting the conformational space available to the dendrimers and enforcing some degree of conformational homogeneity. A more detailed view into dendrimer–substrate interactions reveals the crucial stabilizing role of hydrogen bonds and ion-pairs between sulfonate groups of substrate molecules and the dendrimer residues.

In conclusion, the present work provides fundamental knowledge on how the pH-dependent protonation equilibrium influences the behavior of dendritic systems, yielding new possibilities for the development of applications that take advantage of pH effects in these systems.

5.6 Acknowledgments

We thank Jean-Louis Reymond, João M. Damas, Diogo Vila-Viçosa and Pedro Magalhães for helpful discussions. We acknowledge the financial support from Fundação para a Ciência e Tecnologia, Portugal, through project grant PTDC/QUI-QUI/100416/2008 and R&D Units, UID/CBQ/-04612/2013 and UID/MULTI/00612/2013. LCSF was supported by FCT fellowship SFRH/BD/76085/2011. We also acknowledge the support from EU COST Action CM1102.

Chapter 6

Key results and conclusions

Proliferam confusões sobre o que é um trabalho de pesquisa original. Não é uma reflexão magna sobre o passado e o futuro do universo, nem é uma nova síntese da filosofia ocidental, de Parménides a Popper. É, habitualmente, uma investigação sobre um tema minúsculo e muito especializado, com conclusões modestas e com impacto reduzido. Mas exige muito trabalho original.

— Nuno Crato, in Passeio Aleatório, 2009

Confusions abound regarding what an original research work is. It is not an all-embracing reflection on the past and the future of the universe, nor a synthesis of western philosophy from Parmenides to Popper. It is, usually an investigation about a tinny subject, highly specialized, with modest conclusions and a reduced impact. But it does require a lot of original work.

translation by Luís Filipe

On the first chapter of this thesis we have attempted to provide the reader with an overall understanding of the chemical system of interest, the questions raised by the lack of atomic-level information on such systems and the benefits that could be gained from such knowledge. It is now the time to summarize and contextualize the key contributions of the work we have performed.

The first idea resulting from our studies, and that must be properly emphasized, is that peptide dendrimers are extremely flexible molecules

able to adopt a myriad of conformational states. Overall, and as one would expected, the conformational behavior of these systems is, to a certain degree, conditioned by their topology, but it mostly depends on the amino acid sequences used.

In chapters 3 and 4 we have performed a comprehensive survey into the structural determinants of peptide dendrimers. For that, we combined computational and experimental approaches, and gathered sufficient data to show that a trade-off between electrostatic effects and hydrogen bond formation controls structure acquisition in these systems. Moreover, by selectively changing the dendrimers sequence, and consequently their net charge, we were able to manipulate the exhibited behaviors; dendrimers with an increasingly higher number of charged residues adopt sequentially less compact structures in solution.

In general, the conformational behavior of these systems seems to result mostly from a charge balance. Although same-charge residues in a dendrimer molecule are topologically constrained to retain some proximity, the high conformational flexibility allows them to move apart in order to reduce charge–charge repulsion and increase solvation, causing the molecules to swell.

Also, the inclusion of charged residues, either negative or positive, into a dendrimer sequence, markedly changes the features of its energy landscape. The roughness of the landscapes increases with the decrease in the number of charged residues.

As shown in Chapter 3, the inclusion of branching residues with shorter side-chain, which results in dendrimers that are topologically more symmetrical, does not significantly alter the structural behavior of these systems.

It is interesting to note that our findings agree with the 2013 study of Falkovich et al. [101], obtained from simulations of poly-L-lysine dendrimers (generations one to five). In those dendrimers, the charged groups available were the terminal groups of the lysine residues found at the core and at the outermost periphery of the dendrimers. The authors concluded that those peptide dendrimers are essentially porous open structures where all charged terminal groups are **equally solvent**

exposed. Furthermore, in the aforementioned publication, the authors also noted that "asymmetry of branching does not influence the structural properties of dendrimers significantly" [101], a result that is consistent with the results we have presented in section 3.4.1.

The results presented in chapter 4 of this thesis also corroborate (and are corroborated by) the 2013 study of Ravi et al. [73] where the interaction between a antimicrobial polycationic third-generation peptide dendrimer [with sequence $(Leu)_8(Dap Leu)_4(Dap Phe)_2 Dap LysNH_2$] and models of eukaryotic and prokaryotic cell membranes were studied using molecular dynamics simulations. The charged groups in that peptide dendrimer are found only in the terminal residues (core and periphery). Ravi et al. observed a significant structural plasticity in that peptide dendrimer, which underwent "significant conformational changes in the course of the MD simulation". Before interacting with the membranes, the dendrimer adopted an extended conformation; once it was localized at the membrane–water interface the dendrimer adopted a coiled structure; and upon deeper penetration into the membrane the dendrimer "changed back to an extended conformation similar to that in the aqueous phase" [73]. Moreover, the successful insertion of the dendrimer into the polyanionic membrane of prokaryotes was driven by **electrostatic** interactions, **hydrogen bonds** between the dendrimer and specific lipopolysaccharide functional groups and the dendrimer's ability to adapt its conformation (**flexibility**).

Overall, the conformational patterns described by both Falkovich et al. and Ravi et al. resemble the ones we have observed for highly charged dendrimers.

Also, the conclusions presented are also consistent with the 2009 report of Javor et al. [98] where several peptide dendrimer of generations one to three, containing at most two charged residues, were found to "exist as conformationally flexible molten globules in aqueous solution" with topologically distant residues being found at close geometrical distances. The conformational patterns described by Javor et al. seem to be similar to the ones observed by us for peptide dendrimers containing a small number of charged residues.

To date, no computational study involving these systems had explicitly addressed the protonation equilibrium taking place at different pH values for dendrimers with multiple ionizable sites. In chapter 5 we reported the first constant-pH MD simulations of any dendritic systems, which included simulations of dendrimers of different generations, and also simulations in the presence of other organic molecules that are able to interact with the dendrimers studied. Such simulations allowed us to explore the role of pH and the existence of pH–dependent effects in these systems.

The results show that first and second generation dendrimers are almost structurally unresponsive to pH variations. The protonation curves of the ionizable residues of those dendrimers reflect the dendrimers high flexibility.

For third and fourth generation dendrimers, pH plays a structuring role, with markedly different behaviors being observed due to pH modulation. Protonation–conformation coupling effects influence several intramolecular interactions, which, in turn, modulate the shape and structure at the different pH values. In general, these dendrimers ionizable residues protonate at lower pH values than the ones observed for lower generation dendrimers; interesting protonation profiles could be identified for certain ionizable residues, whose protonation state is deeply related with the conformations adopted by the dendrimers at each particular pH value.

We have also explored the interactions between higher generation dendrimers and the substrate molecules involved in dendrimer-mediated catalysis. Our results provide the first atomic-level glimpse into how the presence of substrates affects the protonation and conformational behavior of peptide dendrimers at the optimum pH for catalysis, namely by restricting the conformational space available to the dendrimers and enforcing some degree of conformational homogeneity. This structuring effect is a consequence of dendrimer–substrate interactions which occur through stabilizing hydrogen bonds and ion–pairs between the substrates sulfonate groups and the dendrimers residues.

The last ten years have made the synthesis of peptide dendrimers

with reasonable molecular sizes (up to the third generation) feasible in an efficient and straightforward way. The initial steps towards bigger and more complex peptide dendrimers have already been undertaken. As synthetic routes evolve, it is conceivable that new applications also appear. Furthermore, most of the applications developed so far are proofs-of-principle of the potential of these systems, unable to rival with the efficiency of biological systems (e.g. enzymes) or industrial–level synthetic alternatives (e.g. antimicrobials). Those systems need to be fine–tuned in order to become of broad use by the scientific community and the results presented here might contribute to these necessary developments and improvements.

Overall, the conclusions obtained in this work are a contribution to the detailed understanding of dendritic systems composed of amino acids, from which I hope others can capitalize on the path toward the establishment of a solid basis for the rational development of novel functional dendritic systems through a judicious choice of topology and amino acid sequence.

Appendix A

Potential Energy Functions

Functional Form of the GROMOS Force Field

The potential energy function in force fields of the GROMOS [114] family takes the form,

$$\begin{aligned} U(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = & \sum_{bonds} \frac{1}{4} K_b [b^2 - b_0^2]^2 \\ & + \sum_{angles} \frac{1}{2} K_\theta [\cos\theta - \cos\theta_0]^2 \\ & + \sum_{dihedrals} K_\varphi [1 + \cos(\delta)\cos(m\varphi)] \\ & + \sum_{impropers} \frac{1}{2} K_\xi [\xi - \xi_0]^2 \\ & + \sum_{pairs(i,j)} \left[\frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^6} \right] \\ & + \sum_{pairs(i,j)} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r} \frac{1}{r_{ij}} . \end{aligned} \tag{A.1}$$

Covalent bond and bond-angle interactions, as well as improper dihedral angles are described using harmonic terms; however, the harmonic dependence is applied to the square of the bond length (leading to a quartic potential) and to the cosine of the bond-angle. Torsional dihedral angles are represented by a trigonometric function. Nonbonded interactions are the sum of van der Waals (modeled by a Lennard-Jones potential) and electrostatic (modeled by a Coulombic term) interactions.

The force constants for bonds, angles, torsional and improper dihedrals are K_b, K_θ, K_φ and K_ξ, respectively. Ideal bond lengths, angles and improper dihedral angles are b_0, θ_0 and ξ_0. For torsional dihedral angles, $\cos(\delta)$ is the phase shift (restricted to 0 or π, i.e., $\cos(\delta) = \pm 1.0$) and m the multiplicity. b, θ, φ and ξ are the actual bond length, angle, and dihedrals defined by the intervening atoms (2, 3 or 4 atoms). The last two terms in the equation (vdW and electrostatic interactions) are described in Chapter 2 (Section 2.1.3).

Functional Form of the AMBER Force Field

The potential energy function in force fields of the AMBER [111] family takes the form,

$$
\begin{aligned}
U(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = & \sum_{bonds} K_b[b - b_0]^2 \\
& + \sum_{angles} K_\theta[\theta - \theta_0]^2 \\
& + \sum_{dihedrals} \sum_{n} \frac{V_n}{2}[1 + \cos(n\varphi - \delta)] \\
& + \sum_{pairs(i,j)} \varepsilon_{ij} \left[\left(\frac{r_0(i,j)}{r_{ij}} \right)^{12} - 2 \left(\frac{r_0(i,j)}{r_{ij}} \right)^{6} \right] \\
& + \sum_{pairs(i,j)} \frac{q_i q_j}{4\pi\varepsilon_0} \frac{1}{r_{ij}} \, .
\end{aligned}
\tag{A.2}
$$

Covalent bond and bond angle interactions are described using harmonic expressions. Torsional dihedral angles are represented by a trigonometric term. Nonbonded interactions are the sum of the van der Waals (modeled by a Lennard-Jones potential) and electrostatic (modeled by a Coulombic term) interactions.

The force constants for bonds and bond-angles are K_b and K_θ, respectively. Equilibrium bond lengths and angles are b_0, θ_0. Regarding dihedral angles, $V_n/2$ is the force constant, n its periodicity and δ the phase offset. b, θ and φ are the actual bond length, angle, and dihedrals defined by the intervening atoms (2, 3 or 4 atoms). The ε_{ij} and $r_0(i,j)$ in the Lennard-Jones term are the vdW well depth and vdW minimum for the atom pair, respectively. The last term of the equation (electrostatic interactions) is described in Chapter 2 (Section 2.1.3).

Appendix B

Supporting information for "Unraveling the Conformational Determinants of Peptide Dendrimers Using Molecular Dynamics Simulations" (Chapter 3)

GROMOS 53A6 parameters for L-ornithine (Orn) and L-2,3-diaminobutanoic acid (Dab)

Parameter macros employed in the fragments parameterization for the GROMOS96 53A6 force field. The values can be found in reference 114. The scheme presented must be interpreted having in mind the GROMACS package configuration. The abbreviations `gb_x`, `ga_x`, `gd_x` and `gi_x`, mean gromos bond (gb), angle (ga), dihedral (gd) and improper (gi); the x accounts for a number that is associated with a certain parameter value in the force field.

Table B.1: GROMOS96 53A6 FF Fragment Parameters.

ORN				
Atoms	N	N	-0.31000	0
	H	H	0.31000	0
	CA	CH1	0.00000	1
	CB	CH2	0.00000	1
	CG	CH2	0.00000	2
	CD	CH2	0.00000	3
	NE	N	-0.31000	3
	HE	H	0.31000	3
	C	C	0.450	4
	O	O	-0.450	4
Bonds	N	H	`gb_2`	
	N	CA	`gb_21`	
	CA	CB	`gb_27`	
	CA	C	`gb_27`	
	CB	CG	`gb_27`	
	CG	CD	`gb_27`	
	CD	NE	`gb_21`	
	NE	HE	`gb_2`	
	C	O	`gb_5`	
	C	+N	`gb_10`	
Angles	-C	N	H	`ga_32`
	-C	N	CA	`ga_31`
	H	N	CA	`ga_18`
	N	CA	CB	`ga_13`
	N	CA	C	`ga_13`

Continued on next page

Angles	CB	CA	C	ga_13	
	CA	CB	CG	ga_15	
	CB	CG	CD	ga_15	
	CG	CD	NE	ga_13	
	CD	NE	HE	ga_18	
	CA	C	O	ga_30	
	CA	C	+N	ga_19	
	O	C	+N	ga_33	
Dihedrals	-CA	-C	N	CA	gd_14
	-C	N	CA	C	gd_39
	N	CA	CB	CG	gd_34
	N	CA	C	+N	gd_40
	CA	CB	CG	CD	gd_34
	CB	CG	CD	NE	gd_34
	CG	CD	NE	HE	gd_29
Impropers	N	-C	CA	H	gi_i
	CA	N	C	CB	gi_2
	C	CA	+N	O	gi_1
DAB					
Atoms	N	N	-0.31000	0	
	N	N	-0.31000	0	
	H	H	0.31000	0	
	CA	CH1	0.00000	1	
	CB	CH2	0.00000	1	
	CG	CH2	0.00000	2	
	ND	N	-0.31000	2	
	HD	H	0.31000	2	
	C	C	0.450	3	
	O	O	-0.450	3	
Bonds	N	H	gb_2		
	N	CA	gb_21		
	CA	CB	gb_27		
	CA	C	gb_27		
	CB	CG	gb_27		
	CG	ND	gb_21		
	ND	HD	gb_2		
	C	O	gb_5		
	C	+N	gb_10		
Angles	-C	N	H	ga_32	
	-C	N	CA	ga_31	
	H	N	CA	ga_18	

Continued on next page

Angles	N	CA	CB	ga_13	
	N	CA	C	ga_13	
	CB	CA	C	ga_13	
	CA	CB	CG	ga_15	
	CB	CG	ND	ga_13	
	CG	ND	HD	ga_18	
	CA	C	O	ga_30	
	CA	C	+N	ga_19	
	O	C	+N	ga_33	
Dihedrals	-CA	C	N	CA	gd_14
	-C	N	CA	C	gd_39
	N	CA	CB	CG	gd_34
	N	CA	C	+N	gd_40
	CA	CB	CG	ND	gd_34
	CB	CG	ND	HD	gd_29
Impropers	N	-C	CA	H	gi_1
	CA	N	C	CB	gi_2
	C	CA	+N	O	gi_1

Computed pK_a values

Poisson–Boltzmann (PB) and Monte-Carlo (MC) calculations were performed with the programs MEAD [194] and PETIT [128, 133], respectively. For each dendrimer the calculations were performed over a set of conformations selected from the final concatenated trajectory at intervals of 0.2 ns considering pH values ranging from -5 to 20 at 0.2 pH units intervals. The atomic charges and radii used in the PB calculations were derived from the GROMOS 53A6 force field. All PB calculations consisted of finite-difference linear PB calculations performed with the program MEAD (version 2.2.0) using a temperature of 300 K, a molecular surface defined with a solvent probe radius of 1.4 Å, and a Stern (ion exclusion) layer of 2.0 Å. The dielectric constants were 80 for solvent and 4 for the peptide dendrimer. The ionic strength used was 0.15 M. A two-step focusing procedure was used, with consecutive grid spacings of 1.0 and 0.25 Å. The MC runs were performed using 10^5 MC cycles, one cycle consisting of sequential state changes over all individual sites and also all pairs of sites with at least one interaction term above 2.0 pK_a units. The PB calculations used the model compounds and calibrated pK_a values according to reference [139].

Table B.2: Computed pK_a values for residues of the NE-series dendrimers. See Figure 3.1 and Table 3.1 in the article for details about residue position.

Dendrimer	Residues	pK_a
NE12	X_1 (Asp)	4.12
	X_2 (Cys)	11.78
	X_3 (Tyr)	11.13
		10.99
	X_6 (Glu)	4.46
		4.49
		4.24
		4.53
	X_8 (Glu)	4.33
		4.29

Continued on next page

NE12	X_8 (Glu)	4.31 4.36 4.34 4.47 4.32 4.36
NE8	X_1 (Asp)	4.00
	X_2 (Cys)	11.61
	X_3 (Tyr)	11.26 10.95
	X_8 (Glu)	4.32 4.25 4.33 4.33 4.29 4.27 4.29 4.30
NE4	X_1 (Asp)	3.93
	X_2 (Cys)	11.67
	X_3 (Tyr)	11.01 10.77
	X_6 (Glu)	4.27 4.13 4.30 4.20
NE0	X_1 (Asp)	3.72
	X_2 (Cys)	10.67
	X_3 (Tyr)	10.82 10.79

2D Energy landscapes for B1- and C1-series

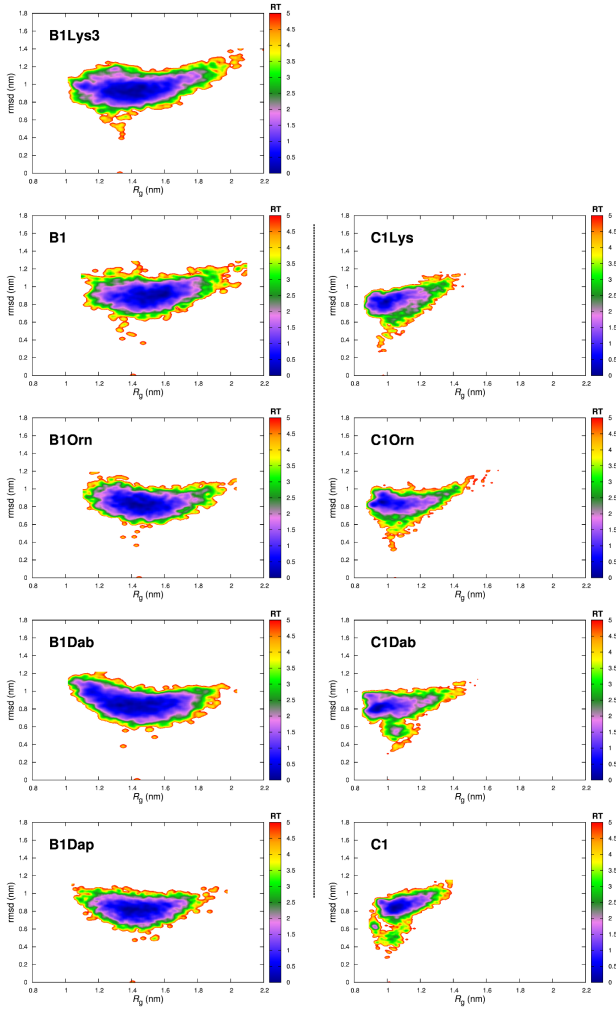


Figure B.1: Energy profiles and corresponding lowest energy conformers for the B1- (left) and C1-series (right). The energy profiles use R_g and rmsd as structural coordinates.

BE-series, amino acid sequences

Table B.3: Composition of the BE-series.

dendrimer	residues at each position[a]											charge	$\langle R_g \rangle$
	X_1	X_2	B_1	X_3	X_4	B_2	X_5	X_6	B_3	X_7	X_8	(e)[b]	(nm)[c]
BE-series													
BE12	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	**Glu**	Dap	Ser	**Glu**	-13	1.44
BE8	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	**Gln**	Dap	Ser	**Glu**	-9	1.30
BE4	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	**Glu**	Dap	Ser	**Gln**	-5	1.14
BE0	Asp	Cys	Dap	Tyr	Amb	Lys	Ala	**Gln**	Dap	Ser	**Gln**	-1	1.02

[a]Notice that BE12 is equivalent to B1. Residue positions in accordance with Figure 3.1. Standard three-letter abbreviations are used for proteinogenic amino acids, Amb for 4-aminomethyl(benzoic) acid and Dap for *L*-2,3-diaminopropanoic acid. The relevant modifications to each series appear underlined. [b] Charge of the simulated dendrimers. [c] Average radius of gyration as computed from the simulations.

BE-series, dendrimer compactness

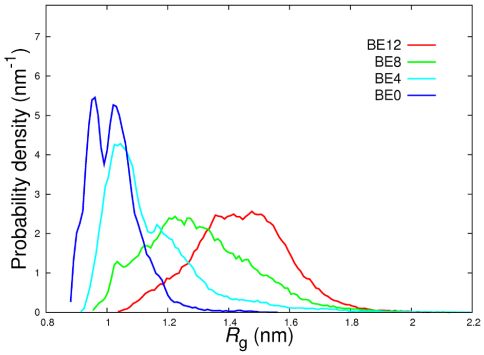


Figure B.2: Radius of gyration probability density histograms for the BE-series.

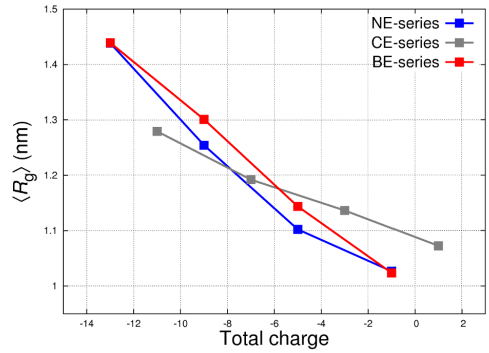


Figure B.3: Dendrimer total charge versus average radius of gyration for the NE- and CE-series, and also for the BE-series.

BE-series, energy landscapes

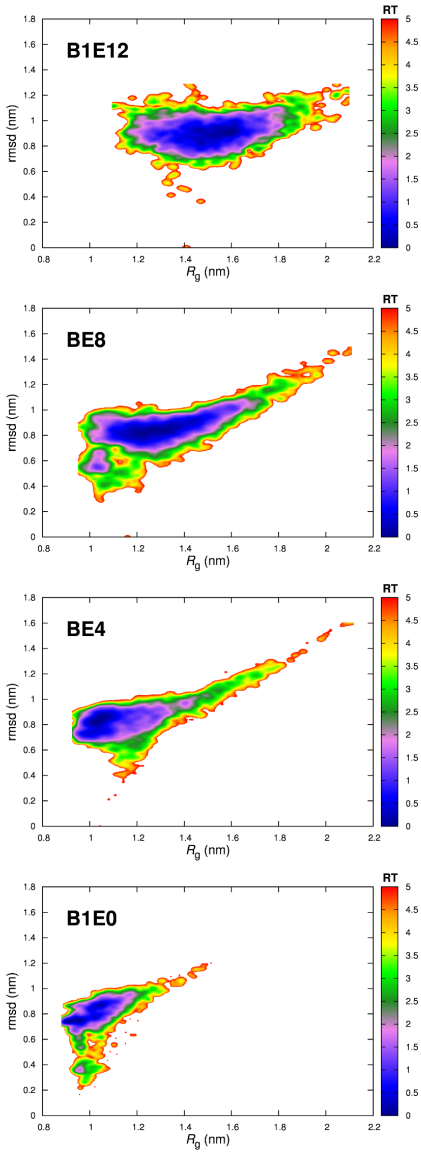


Figure B.4: Energy profiles for the BE-series. The energy profiles use R_g and rmsd as structural coordinates.

Comparison between different electrostatic treatments

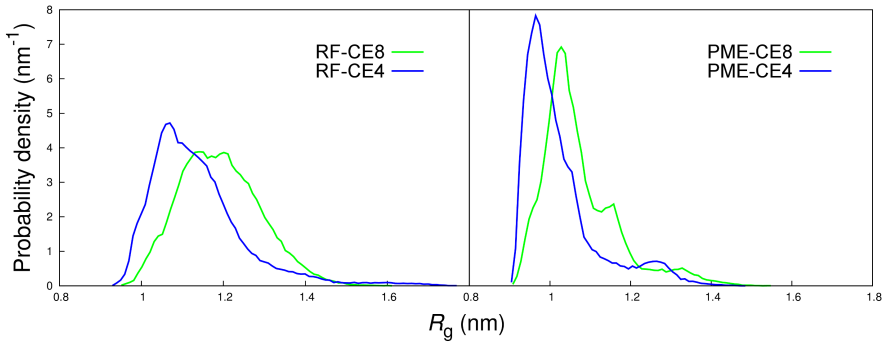


Figure B.5: Radius of gyration probability density histograms for the CE8 and CE4 dendrimers, using the reaction field (RF) and particle-mesh Ewald (PME) methods to treat long-range electrostatic interactions. PME simulations used a real-space cutoff of 0.9 nm and a 0.16 nm grid spacing; other conditions as described in section 3.3.1 of Chapter 3.

The PME treatment leads to a higher compactness than the reaction field, but the same qualitative effect is observed, with the increase of charge leading to an increase of radius of gyration. Several studies indicate that Ewald summation methods may artificially enhance structural stability and hinder conformational sampling [e.g., see refs. 7-12 in [238] and refs. 63-69 in [156], which may explain this overall compaction. Additionally, the fact that both the solute and water models were parameterized using the reaction-field approach may also contribute to explain this behavior.

Average R_g versus average number of hydrogen bonds

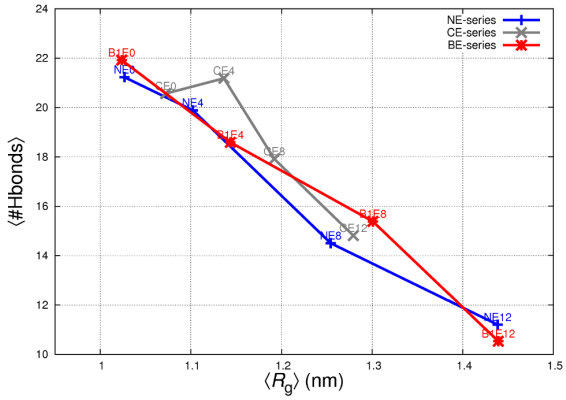


Figure B.6: Average radius of gyration values for the different dendrimer series as function of the average number of hydrogen bonds established.

CE- and BE-series hydrogen bond networks

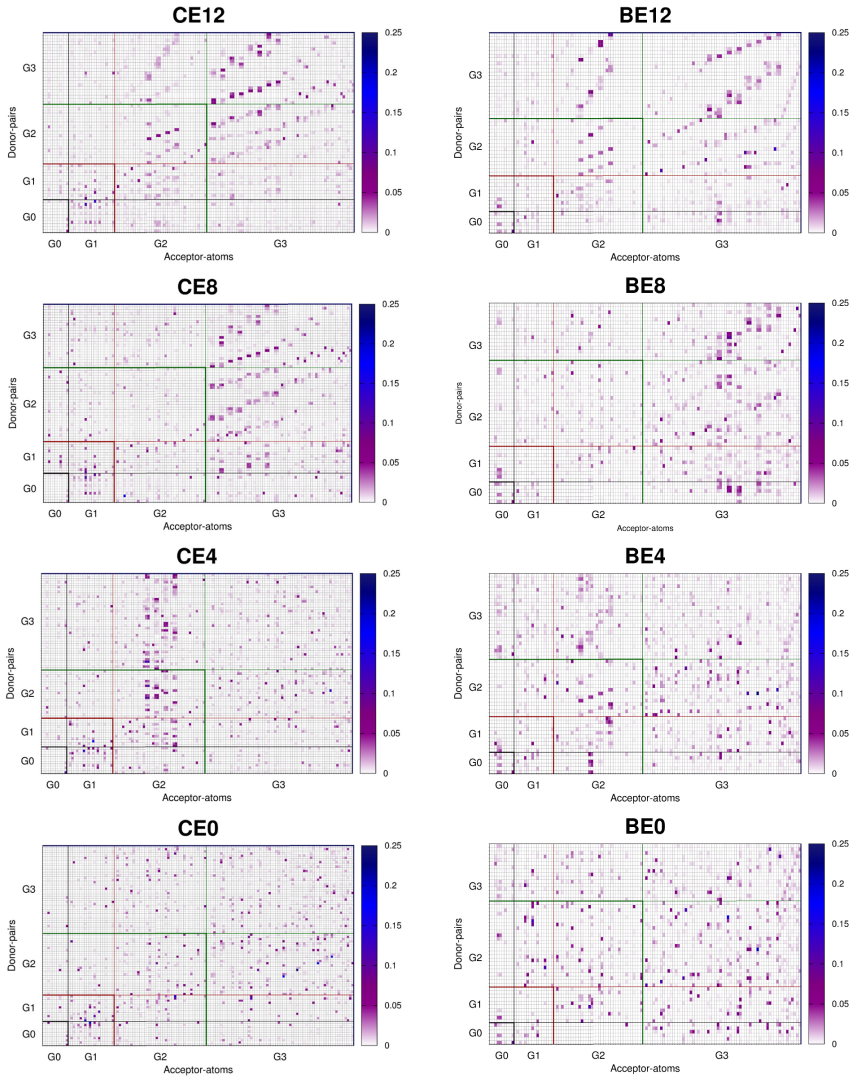


Figure B.7: Hydrogen bond matrices for the CE-series and BE-series dendrimers.

CE- and BE-series hydrogen bonds categorized

Table B.4: CE-series. Average number of hydrogen bonds formed between donor-acceptor atoms belonging to: main-chain, side-chains of neutral residues (neutral), side-chains of negatively charged residues (negative) and side-chains of positively charged residues (positive).

Dendrimer	Donors	Acceptors			
		Main-Chain	Neutral	Negative	Positive
CE12	Main-Chain	**4.03**	1.17	4.81	0.65
	Neutral	0.75	0.23	2.00	0.09
	Positive	0.58	0.18	0.27	0.004
CE8	Main-Chain	**6.27**	1.97	3.52	0.83
	Neutral	1.25	0.53	2.29	0.15
	Positive	0.51	0.22	0.29	0.08
CE4	Main-Chain	**6.72**	2.84	3.86	0.80
	Neutral	2.29	1.08	2.15	0.18
	Positive	0.70	0.32	0.19	0.05
CE0	Main-Chain	**8.92**	4.72		0.56
	Neutral	2.93	2.43		0.23
	Positive	0.46	0.27		0.03

Table B.5: BE-series. Average number of hydrogen bonds formed between donor-acceptor atoms belonging to: main-chain, side-chains of neutral residues (neutral) and side-chains of negatively charged residues (negative).

Dendrimer	Donors	Acceptors		
		Main-Chain	Neutral	Negative
BE12	Main-Chain	**3.67**	0.64	4.42
	Neutral	0.54	0.21	1.06
BE8	Main-Chain	**5.01**	1.54	5.28
	Neutral	1.24	0.55	1.76
BE4	Main-Chain	**7.39**	2.97	3.00
	Neutral	2.52	1.02	1.69
BE0	Main-Chain	**10.16**	4.70	0.38
	Neutral	3.89	2.25	0.53

Appendix C

Supporting information for "Exploring the structural properties of positively charged peptide dendrimers" (Chapter 4)

Material and Reagents

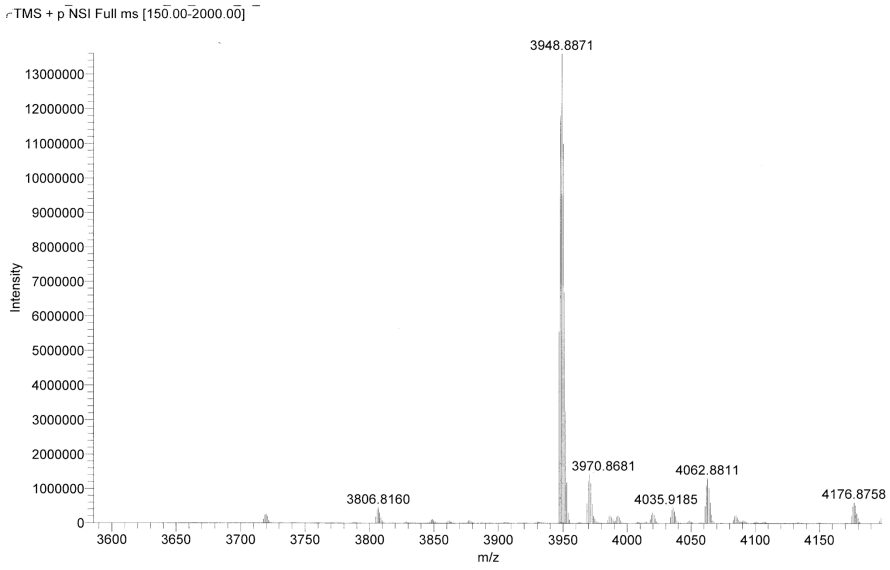
All reagents were purchased in the highest quality available from Aldrich, Fluka or Acros Organics. PyBOP, Fmoc-protected amino acids and their derivatives were purchased from Advanced ChemTech (USA), Novabiochem (Switzerland), IRIS Biotech (Germany) PolyPeptide (France) and GL BioChem (China). The following amino acid derivatives were used: Fmoc-Asp(OtBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Dap(Fmoc)-OH, Fmoc-Tyr(tBu)-OH, 4-(Fmoc-aminomethyl)benzoic acid, Fmoc-Ala-OH, Fmoc-Dab(Boc)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ser(tBu)-OH. TentaGel S RAM resin (loading: $0.26\,\mathrm{mmol\,g^{-1}}$) was purchased from Rapp Polymere (Germany). All the solvents used were p.a. quality and distilled prior to use.

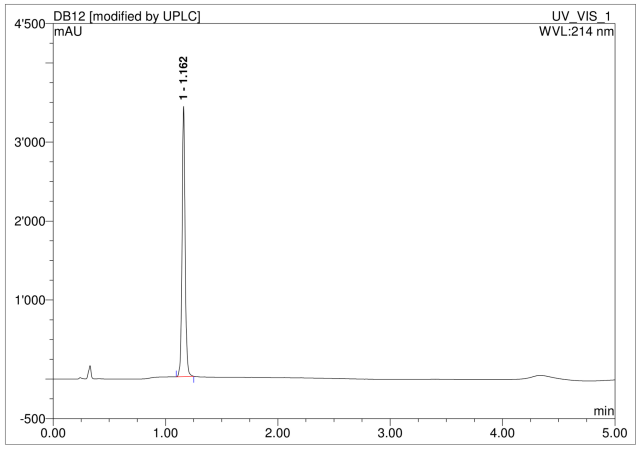
Preparative RP-HPLC (flow rate $80\,\mathrm{mL\,min^{-1}}$) was performed on Waters Delta Prep LC4000 Preparative Chromatography System with a Waters Prepak Cartridge with Delta-Pak C18, 15 µm, 300 Å pore size as column and Waters 486 tunable absorbance detector (λ= 214 nm).

Analytical RP-UHPLC was performed in a Dionex ULTIMATE 3000 Rapid Separation LC System (ULTIMATE 3000RS diode array detector) using a Dionex Acclaim RSLC 120 C18 column (3.0×50 mm, 120 Å, 2.2 µm, flow rate $1.2\,\mathrm{mL\,min^{-1}}$). Compounds were detected by UV absorption at 214 nm. Data recording and processing was done with Dionex Chromeleon Management System Version 6.80.

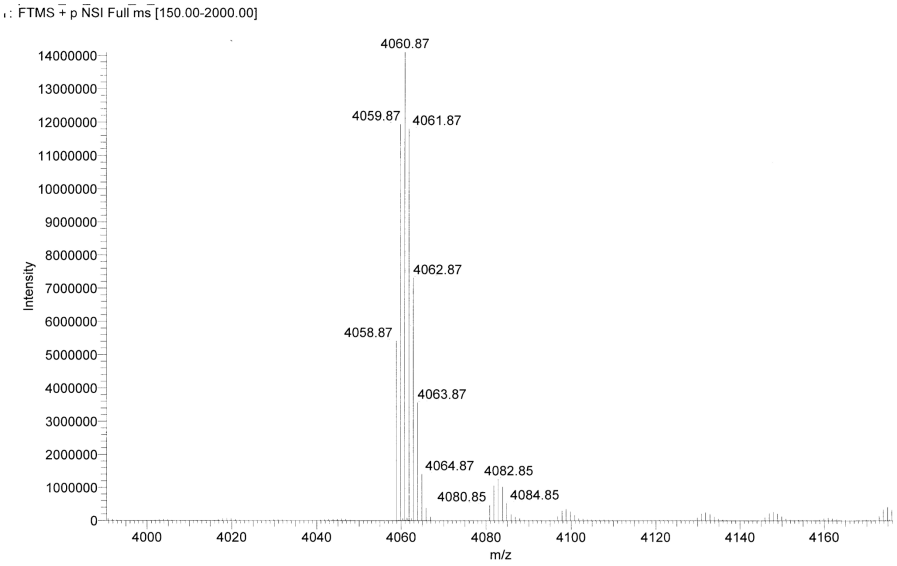
All RP-HPLC were done using HPLC-grade acetonitrile and miliQ-deionized water. Eluents for all systems were: (A) H_2O with 0.1% TFA; (D) H_2O/MeCN (40:60) with 0.1% TFA. The MS spectra and PGSE-NMR measurements were provided by the Mass Spectrometry and NMR services of the Department of Chemistry and Biochemistry at the University of Berne.

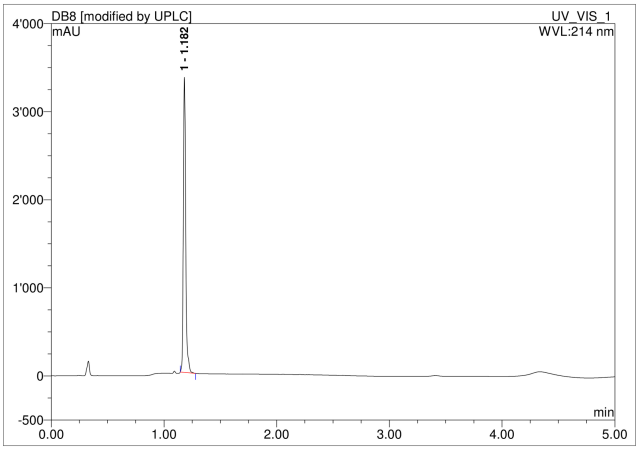
P12 (Ac**Dab**Ser)$_8$(*Dap***Dab**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$. From TentaGel S RAM resin (500 mg, 0.26 mmol g^{-1}), dendrimer **P12** was obtained as a white foamy solid after preparative RP-HPLC purification (16.9 mg, 2.4%). Analytical RP-HPLC: tR = 1.16 min (A/D = 100/0 to 0/100, 1.2 mL min^{-1}, λ= 214 nm). MS (ESI+) calc. for $C_{162}H_{259}N_{57}O_{57}S$ $[M+H]^+$: 3947.9, obsd: 3948.9.



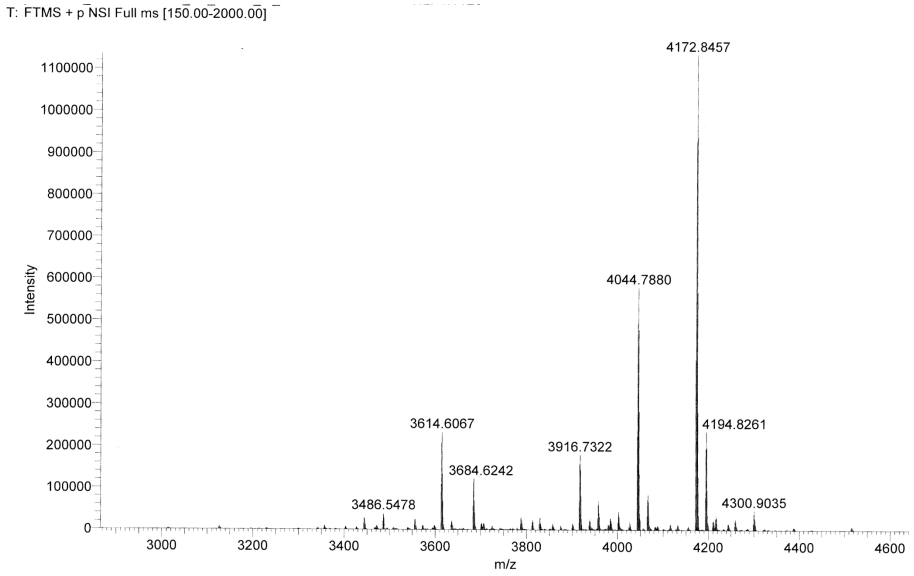


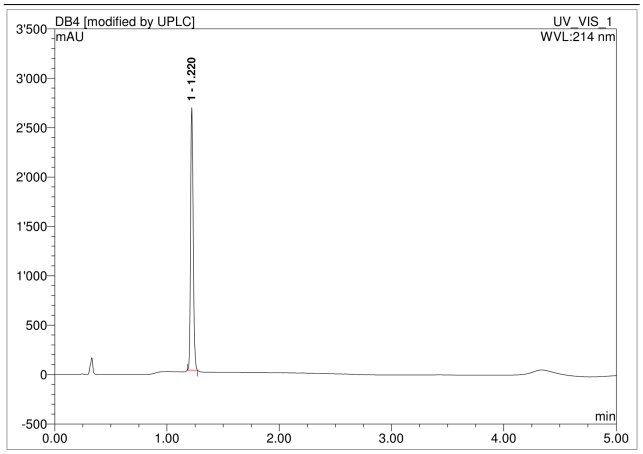
P8 (Ac**Dab**Ser)$_8$(*Dap***Gln**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$. From TentaGel S RAM resin (500 mg, 0.26 mmol g^{-1}), dendrimer **P8** was obtained as a white foamy solid after preparative RP-HPLC purification (35.1 mg, 5.3%). Analytical RP-HPLC: tR = 1.18 min (A/D = 100/0 to 0/100, 1.2 mL min^{-1}, λ= 214 nm). MS (ESI+) calc. for $C_{166}H_{259}N_{57}O_{61}S$ $[M+H]^+$: 4059.9, obsd: 4060.9.





P4 (Ac**Gln**Ser)$_8$(*Dap***Dab**Ala)$_4$(*Dap*AmbTyr)$_2$*Dap*CysAspNH$_2$. From TentaGel S RAM resin (500 mg, 0.26 mmol g^{-1}), dendrimer **P4** was obtained as a white foamy solid after preparative RP-HPLC purification (50.3 mg, 8.2%). Analytical RP-HPLC: tR = 1.22 min (A/D = 100/0 to 0/100, 1.2 mL min^{-1}, λ= 214 nm). MS (ESI+) calc. for $C_{170}H_{259}N_{57}O_{65}S$ $[M+H]^+$: 4171.8, obsd: 4172.8.





Partial Atomic Charges

Partial atomic charges for 4-aminomethyl(benzoic) acid (Amb) and the branching *L*-2,3-diaminopropanoic acid (*Dap*) were derived in accordance with the methodology originally employed in the Amber ff99SB force field [112, 198, 239]. Chemically blocked versions of each residue were constructed: Ac-Amb-NME and $(Ac)_2$-Dap-NME; where Ac are acetyl caps and NME is N-methyl. Each amino acid was represented in two conformations, specifically the extended (C5/β-sheet) and α-helical conformations [239]. All peptide bonds were in the *trans* conformation.

For each of those two conformations of each aminoacid, a quantum-mechanics geometry optimization was performed with the GAMESS-US software [240] using the Hartree-Fock method with the 6-31G* basis set. Electrostatic potentials were then calculated at the same level for the minimized geometries. Finally, the restrained electrostatic potential (RESP) [241] approach was used to derive the partial atomic charges in a two-stage multiconformational fitting, applying restraint potentials of 0.0005 (applied to all heavy atoms) and 0.0010 (methyl and methylene groups) in each stage, respectively.

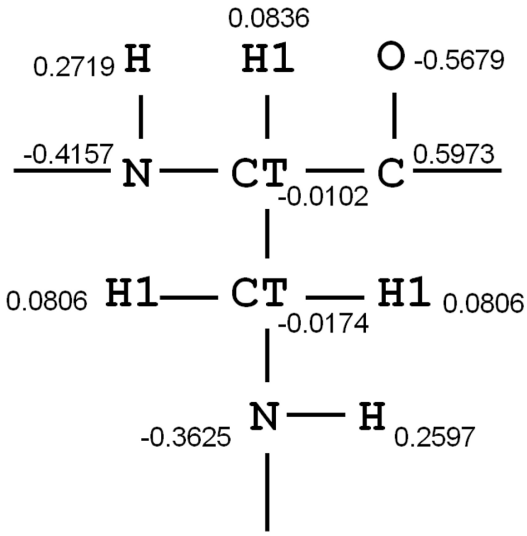


Figure C.1: Atom types and atomic partial charges for *L*-2,3-diaminopropanoic acid (Dap)

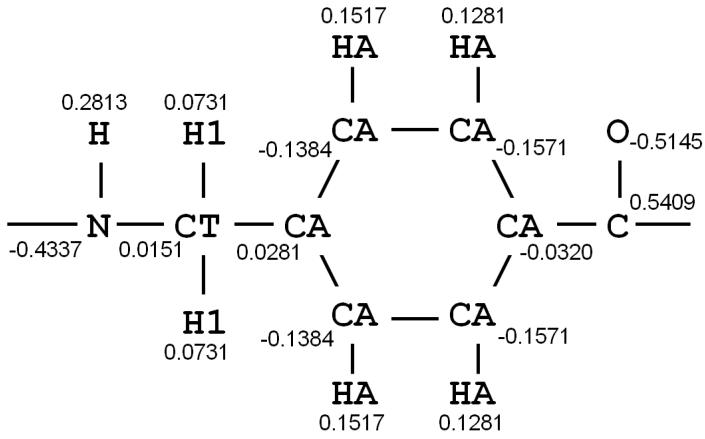


Figure C.2: Atom types and atomic partial charges for 4-aminomethyl(benzoic) acid (Amb)

Relative Solvent Accessibilities

Relative solvent accessibilities were defined as a residue solvent accessibility (ASA) normalized by a reference maximum ASA value for that residue [211]. Solvent accessible surface areas for each residue were computed using the method described in reference 242 with a 1.4 Å spherical probe. RAS values were calculated for each residue and the averaging was done over all the snapshots from the MD simulations.

The normalization factors for the different amino acids were derived by evaluating the surface area around each residue (**X**) when placed in a Ac-Gly-**X**-Gly-NH_2 tripeptide. For the branching residue, a (Ac-Gly)$_2$*Dap*-Gly-NH_2 branched peptide was used. For each peptide, different φ/ψ combinations of extended conformations suggested in the literature were tested [211, 243, 244]. The highest ASA value observed in the different conformations was taken as the reference maximum ASA (Table C.1).

Table C.1: Normalization factors used in Relative Solvent Accessibility calculations

Residue	**Maximum ASA (Å^2)**
Ala	148
Amb	241
Asp	182
Bap	105
Cys	165
Dab	183
Gln	223
Ser	155
Tyr	268

Appendix D

Supporting information for "Structuring peptide dendrimers through pH modulation and substrate binding" (Chapter 5)

Replicate sampling convergence

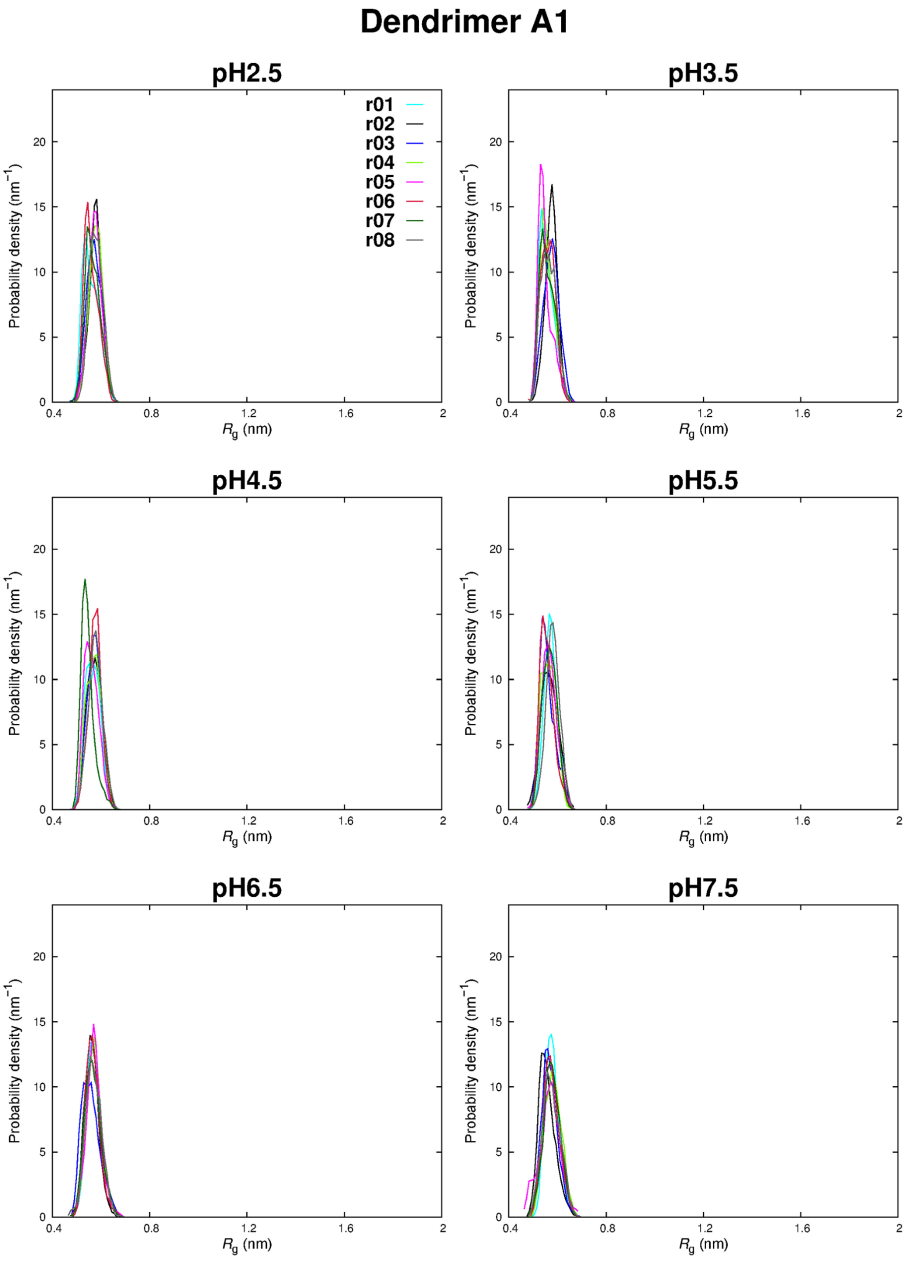


Figure D.1: Dendrimer A1. Radius of gyration probability density histograms for the different replicates at the pH values studies.

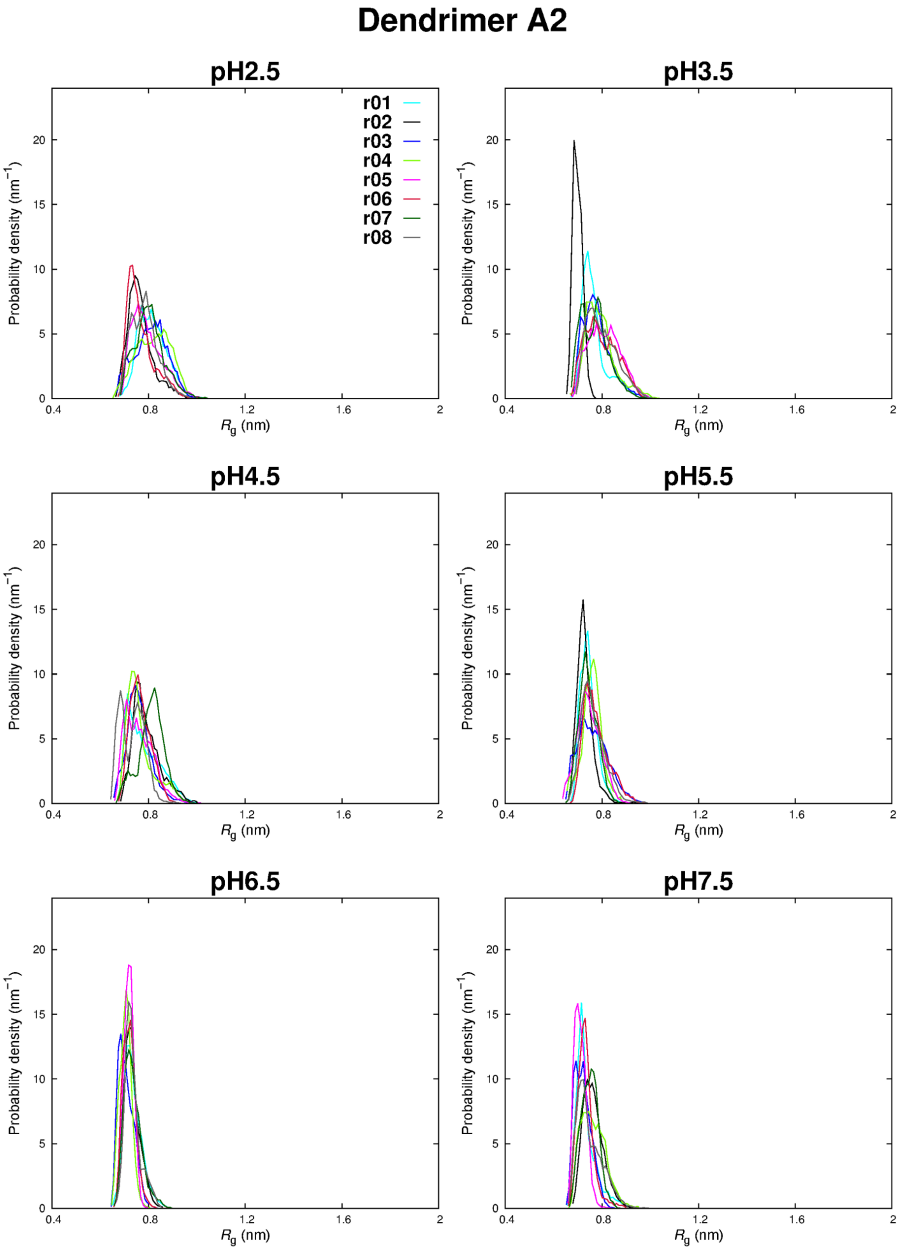


Figure D.2: Dendrimer A2. Radius of gyration probability density histograms for the different replicates at the pH values studies.

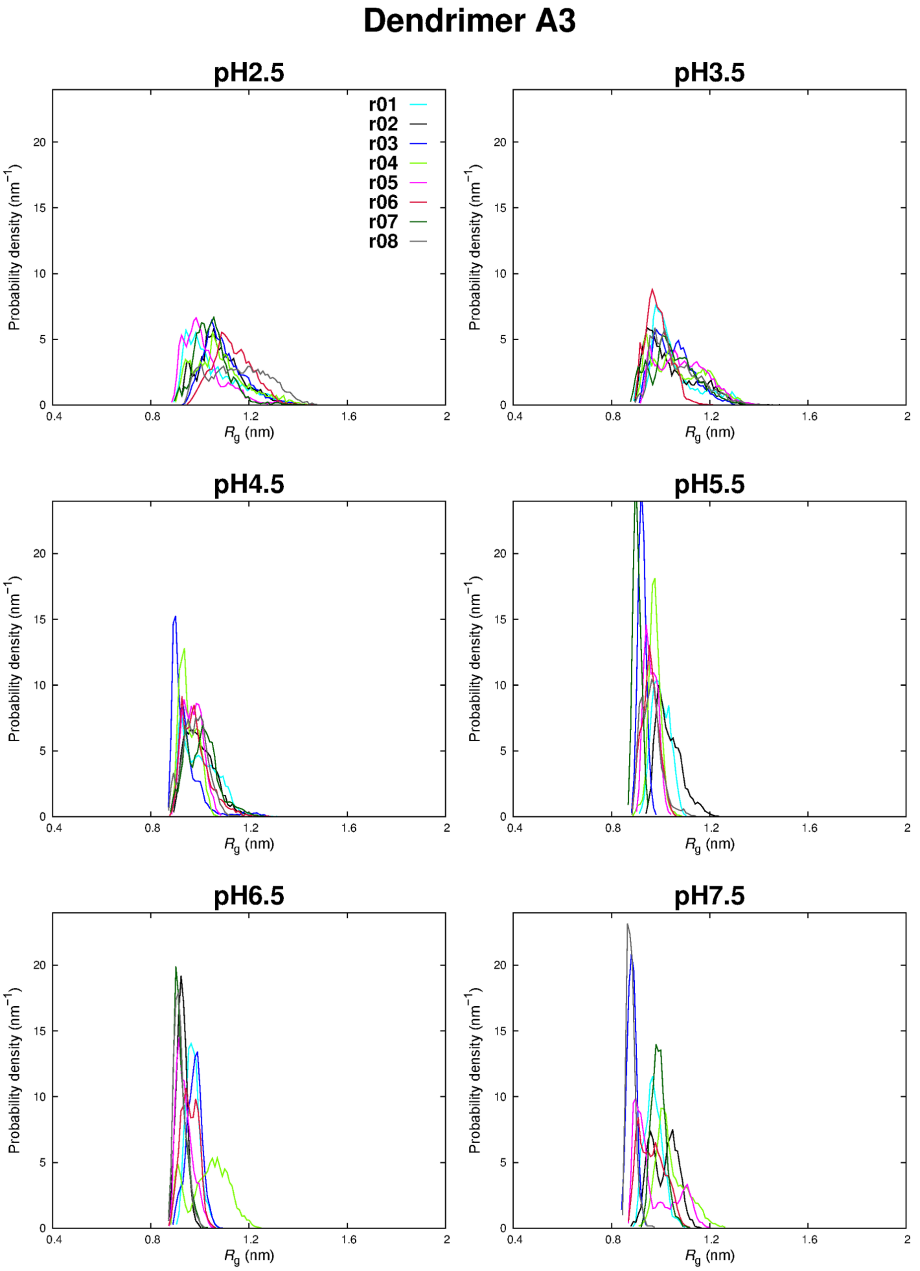


Figure D.3: Dendrimer A3. Radius of gyration probability density histograms for the different replicates at the pH values studies.

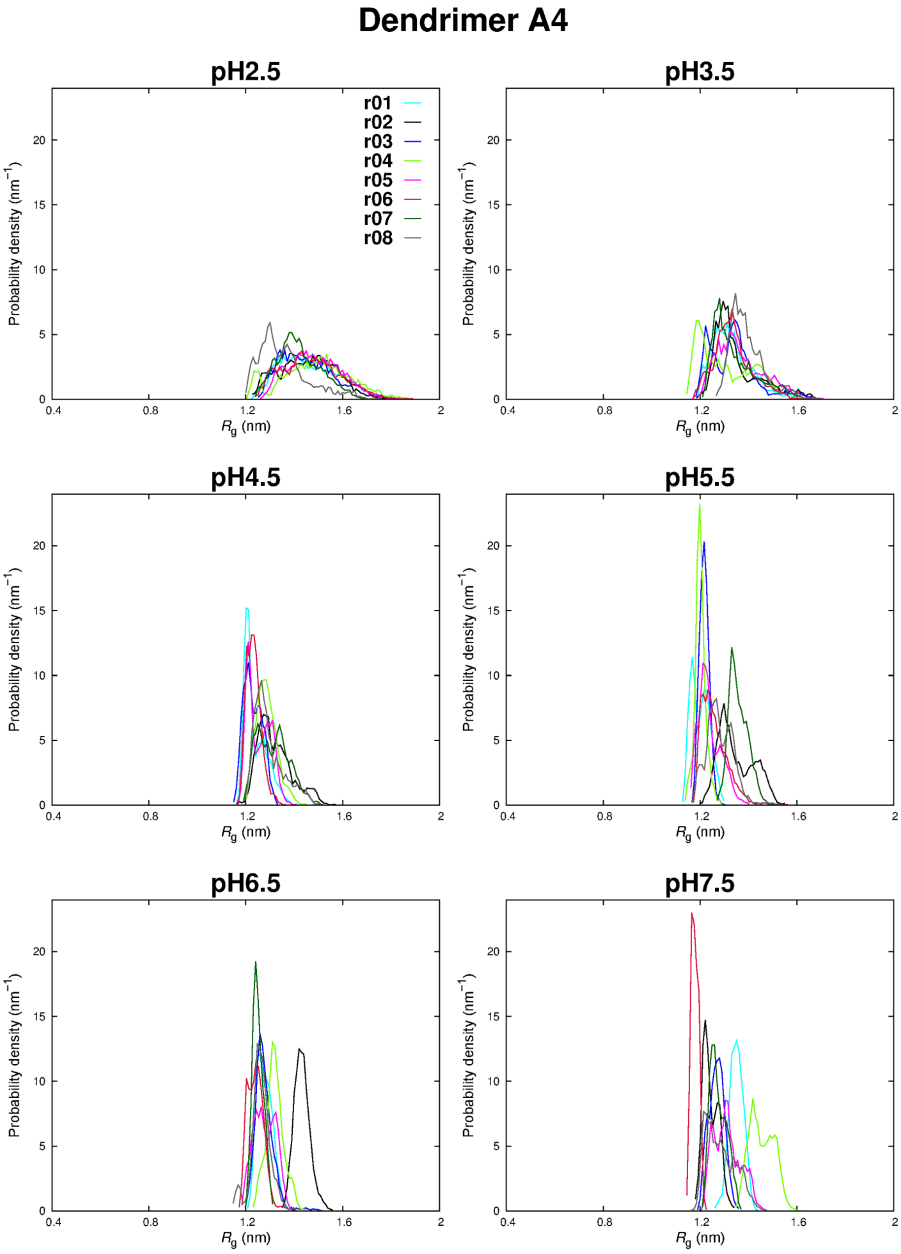


Figure D.4: Dendrimer A4. Radius of gyration probability density histograms for the different replicates at the pH values studies.

Parameterization of substrate BPTS

Partial atomic charges were determined using a procedure described elsewhere [245]. The molecule geometry was optimized with Gaussian 03 at the HF/6-31G(d,p) level, and the resulting electrostatic potential was fitted to atomic coordinates with RESP [241]. The final atomic charges were slightly adjusted to be consistent with the charge groups presented in Figure D.5.

The bonded and van der Waals parameters for 8-butyryloxypyrene-1,3,6-trisulfonate were derived from the GROMOS 54A7 forcefield [115]. The parameter macros presented in tables D.1 to D.6 must be interpreted having in mind the GROMACS package configuration. The abbreviations gb *x*, ga *x*, gd *x* and gi *x*, mean gromos bond (gb), angle (ga), dihedral (gd) and improper (gi); the x accounts for a number that is associated with a certain parameter value in the force field.

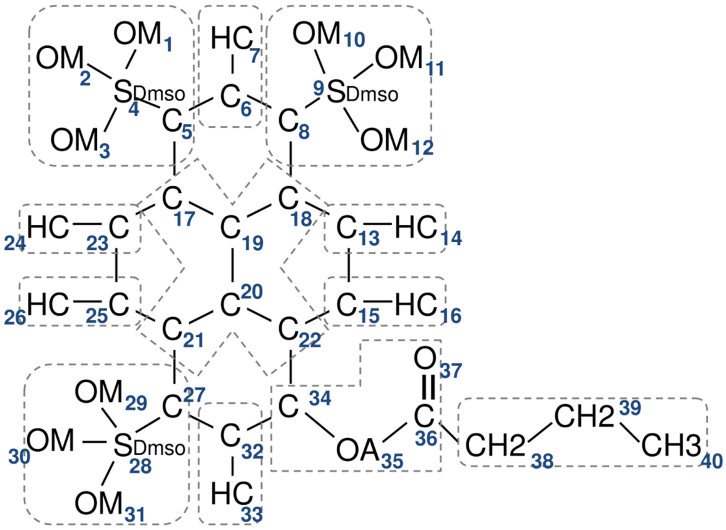


Figure D.5: Atom numbering and atom types for 8-butyryloxypyrene-1,3,6-trisulfonate (**substrate BPTS)**. Dashed lines delimit the charged-groups considered.

Table D.1: Atomic partial charges of BPTS

Atom	Charge	Atom	Charge	Atom	Charge
1	-0.610	**15**	-0.140	**29**	-0.610
2	-0.610	**16**	0.140	**30**	-0.610
3	-0.610	**17**	0.000	**31**	-0.610
4	0.910	**18**	0.000	**32**	-0.140
5	-0.080	**19**	0.000	**33**	0.140
6	-0.140	**20**	0.000	**34**	0.340
7	0.140	**21**	0.000	**35**	-0.400
8	-0.080	**22**	0.000	**36**	0.580
9	0.910	**23**	-0.140	**37**	-0.520
10	-0.610	**24**	0.140	**38**	0.000
11	-0.610	**25**	-0.140	**39**	0.000
12	-0.610	**26**	0.140	**40**	0.000
13	-0.140	**27**	-0.080		
14	0.140	**28**	0.910		

Table D.2: Bond types of BPTS

Atom(i)	Atom(j)	Bond type	Atom(i)	Atom(j)	Bond type
1	**4**	gb 25	**20**	**22**	gb 16
2	**4**	gb 25	**21**	**20**	gb 16
3	**4**	gb 25	**21**	**27**	gb 16
4	**5**	gb 32	**22**	**34**	gb 16
5	**17**	gb 16	**23**	**24**	gb 3
5	**6**	gb 16	**23**	**25**	gb 16
6	**7**	gb 3	**25**	**26**	gb 3
6	**8**	gb 16	**25**	**21**	gb 16
8	**9**	gb 32	**27**	**32**	gb 16
8	**18**	gb 16	**27**	**28**	gb 32
9	**10**	gb 25	**28**	**29**	gb 25
9	**11**	gb 25	**28**	**30**	gb 25
9	**12**	gb 25	**28**	**31**	gb 25
13	**14**	gb 3	**32**	**33**	gb 3
13	**15**	gb 16	**32**	**34**	gb 16
15	**16**	gb 3	**34**	**35**	gb 18
15	**22**	gb 16	**35**	**36**	gb 13
17	**23**	gb 16	**36**	**37**	gb 5
18	**13**	gb 16	**36**	**38**	gb 27
18	**19**	gb 16	**38**	**39**	gb 27
19	**20**	gb 16	**39**	**40**	gb 27
19	**17**	gb 16			

Table D.3: Angle types of BPTS

Atom(i)	Atom(j)	Atom(k)	Angle type	Atom(i)	Atom(j)	Atom(k)	Angle type
1	**4**	**2**	ga 14	**17**	**23**	**24**	ga 25
1	**4**	**3**	ga 14	**17**	**23**	**25**	ga 27
1	**4**	**5**	ga 46	**18**	**19**	**20**	ga 27
2	**4**	**3**	ga 29	**19**	**17**	**23**	ga 27
2	**4**	**5**	ga 46	**19**	**20**	**21**	ga 27
3	**4**	**5**	ga 46	**19**	**20**	**22**	ga 27
4	**5**	**6**	ga 35	**20**	**21**	**25**	ga 27
4	**5**	**17**	ga 35	**20**	**21**	**27**	ga 27
5	**6**	**7**	ga 25	**20**	**22**	**34**	ga 27
5	**6**	**8**	ga 27	**21**	**20**	**22**	ga 27
5	**17**	**19**	ga 27	**21**	**25**	**23**	ga 27
5	**17**	**23**	ga 35	**21**	**25**	**26**	ga 25
6	**5**	**17**	ga 27	**21**	**27**	**28**	ga 35
6	**8**	**9**	ga 35	**21**	**27**	**32**	ga 27
6	**8**	**18**	ga 27	**22**	**34**	**32**	ga 27
7	**6**	**8**	ga 25	**22**	**34**	**35**	ga 19
8	**9**	**10**	ga 46	**23**	**25**	**26**	ga 25
8	**9**	**11**	ga 46	**24**	**23**	**25**	ga 25
8	**9**	**12**	ga 46	**25**	**21**	**27**	ga 35
8	**18**	**13**	ga 35	**27**	**28**	**29**	ga 46
8	**18**	**19**	ga 27	**27**	**28**	**30**	ga 46
9	**8**	**18**	ga 35	**27**	**28**	**31**	ga 46
10	**9**	**11**	ga 14	**27**	**32**	**33**	ga 25
10	**9**	**12**	ga 14	**27**	**32**	**34**	ga 27
11	**9**	**12**	ga 29	**28**	**27**	**32**	ga 35
13	**15**	**16**	ga 25	**29**	**28**	**30**	ga 14
13	**15**	**22**	ga 27	**29**	**28**	**31**	ga 14
13	**18**	**19**	ga 27	**30**	**28**	**31**	ga 29
14	**13**	**15**	ga 25	**32**	**34**	**35**	ga 19
14	**13**	**18**	ga 25	**33**	**32**	**34**	ga 25
15	**13**	**18**	ga 27	**34**	**35**	**36**	ga 12
15	**22**	**20**	ga 27	**35**	**36**	**37**	ga 33
15	**22**	**34**	ga 35	**35**	**36**	**38**	ga 19
16	**15**	**22**	ga 25	**36**	**38**	**39**	ga 15
17	**18**	**19**	ga 27	**37**	**36**	**38**	ga 30
17	**19**	**20**	ga 27	**38**	**39**	**40**	ga 15

Table D.4: Dihedrals of BPTS

i	*j*	*k*	*n*	**type**	**dihedral**	*i*	*j*	*k*	*n*	**type**	**dihedral**
6	**5**	**4**	**3**	1	gd 40	**19**	**20**	**22**	**15**	2	gi 1
6	**8**	**9**	**12**	1	gd 40	**20**	**19**	**18**	**13**	2	gi 1
22	**34**	**35**	**36**	1	gd 17	**20**	**21**	**19**	**22**	2	gi 1
32	**27**	**28**	**29**	1	gd 40	**20**	**22**	**15**	**13**	2	gi 1
34	**35**	**36**	**38**	1	gd 11	**21**	**20**	**27**	**25**	2	gi 1
35	**36**	**38**	**39**	1	gd 40	**22**	**15**	**13**	**18**	2	gi 1
36	**38**	**39**	**40**	1	gd 34	**22**	**15**	**34**	**20**	2	gi 1
5	**6**	**4**	**17**	2	gi 1	**22**	**20**	**19**	**18**	2	gi 1
5	**6**	**8**	**18**	2	gi 1	**23**	**17**	**19**	**20**	2	gi 1
5	**17**	**19**	**18**	2	gi 1	**23**	**17**	**24**	**25**	2	gi 1
8	**6**	**9**	**18**	2	gi 1	**23**	**25**	**21**	**20**	2	gi 1
13	**18**	**14**	**15**	2	gi 1	**25**	**21**	**20**	**19**	2	gi 1
15	**13**	**16**	**22**	2	gi 1	**25**	**21**	**26**	**23**	2	gi 1
17	**5**	**6**	**8**	2	gi 1	**25**	**23**	**17**	**19**	2	gi 1
17	**19**	**5**	**23**	2	gi 1	**27**	**21**	**20**	**22**	2	gi 1
17	**19**	**18**	**8**	2	gi 1	**27**	**21**	**28**	**32**	2	gi 1
17	**19**	**20**	**21**	2	gi 1	**27**	**32**	**34**	**22**	2	gi 1
17	**23**	**25**	**21**	2	gi 1	**32**	**27**	**21**	**20**	2	gi 1
18	**13**	**8**	**19**	2	gi 1	**32**	**34**	**22**	**20**	2	gi 1
19	**17**	**5**	**6**	2	gi 1	**34**	**22**	**20**	**21**	2	gi 1
19	**17**	**20**	**18**	2	gi 1	**34**	**22**	**35**	**32**	2	gi 1
19	**18**	**8**	**6**	2	gi 1	**34**	**32**	**27**	**21**	2	gi 1
19	**18**	**13**	**15**	2	gi 1	**36**	**35**	**37**	**38**	2	gi 1

Table D.5: Pairs in BPTS

Atom(i)	Atom(j)	type	Atom(i)	Atom(j)	type
1	**6**	1	**21**	**29**	1
1	**17**	1	**21**	**30**	1
2	**6**	1	**21**	**31**	1
2	**17**	1	**22**	**36**	1
3	**6**	1	**29**	**32**	1
3	**17**	1	**30**	**32**	1
6	**10**	1	**31**	**32**	1
6	**11**	1	**32**	**36**	1
6	**12**	1	**34**	**37**	1
10	**18**	1	**34**	**38**	1
11	**18**	1	**36**	**40**	1
12	**18**	1	**37**	**39**	1

Table D.6: Exclusions in BPTS

Atom(i)	Atom(j)	Atom(i)	Atom(j)	Atom(i)	Atom(j)
4	7	13	20	20	23
4	8	13	34	20	26
4	19	14	16	20	28
4	23	14	19	20	32
5	9	14	22	20	35
5	18	15	19	21	24
5	20	15	21	21	34
5	24	15	32	22	25
5	25	15	35	22	27
6	13	16	18	22	33
6	19	16	20	23	27
6	23	16	34	24	26
7	9	17	21	25	28
7	17	17	22	25	32
7	18	17	26	25	33
8	14	18	21	26	27
8	15	18	22	27	35
8	17	18	23	28	33
8	20	19	24	28	34
9	13	19	25	33	35
9	19	19	27	35	39
13	17	19	34		

Dendrimer-substrate equilibration times

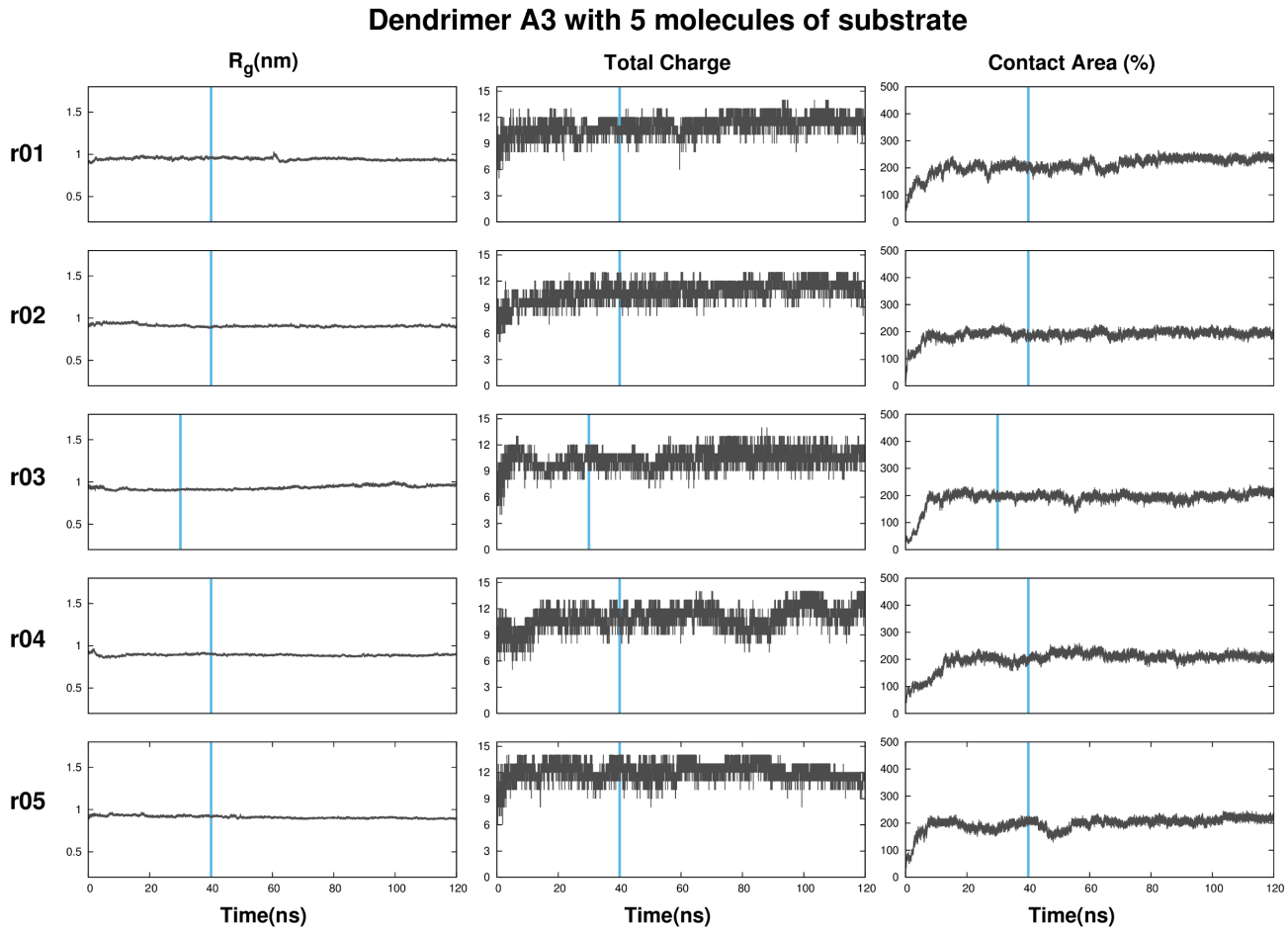


Figure D.6: Replicates time series of: Radius of gyration (*left*), total charge (*middle*), and sum of all the substrates contact areas (*right*). Dendrimer A3 with five substrates at pH5.5. Vertical blue lines signal the equilibration time chosen.

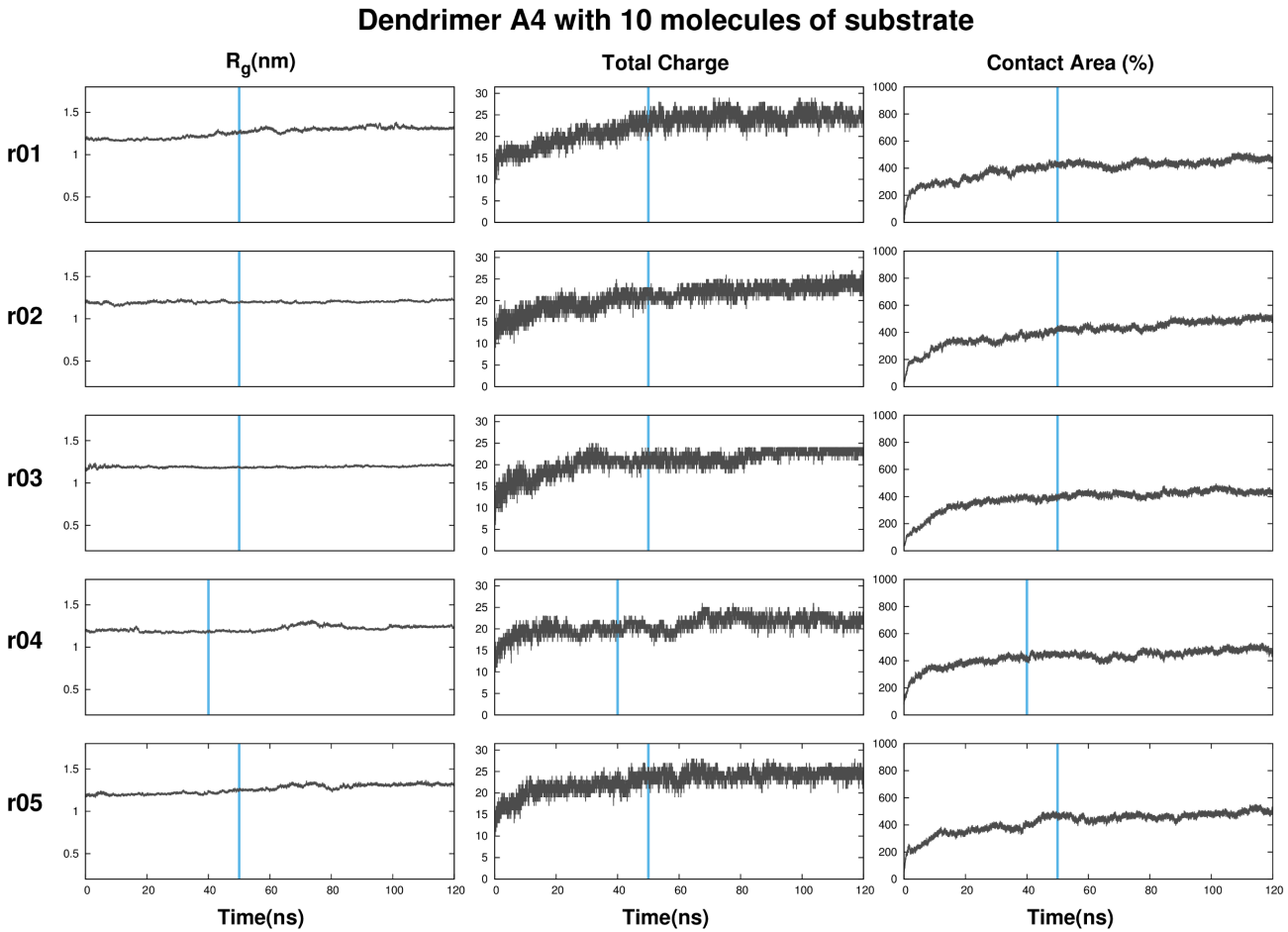


Figure D.7: Replicates time series of: Radius of gyration (*left*), total charge (*middle*), and sum of all the substrates contact areas (*right*). Dendrimer A4 with ten substrates at pH5.5. Vertical blue lines signal the equilibration time chosen.

Dendrimer–substrate replicate sampling convergence

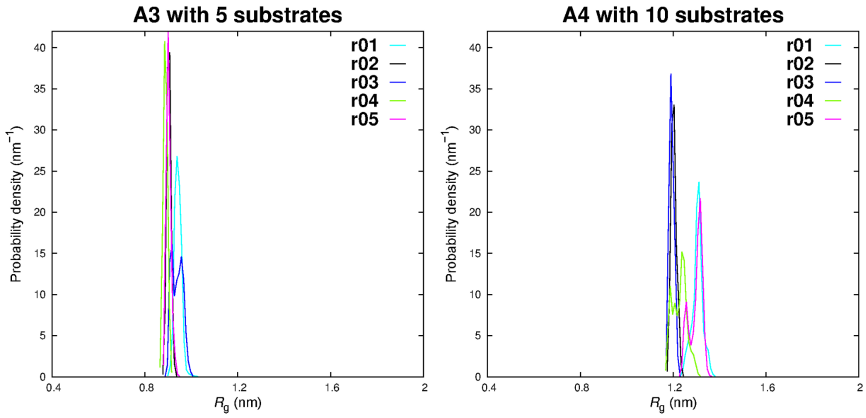


Figure D.8: Radius of gyration probability density histograms for the different replicates. Dendrimer A3 (left) and dendrimer A4 (right).

X-H···π test

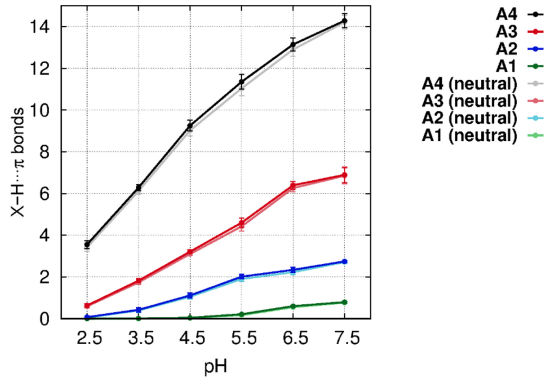


Figure D.9: We tested the robustness of the criteria employed for X-H···π bonds by computing X-H···π interactions considering all histidine residues in the calculations, regardless of their charged form, and performing the same calculations while considering only neutral histidines (neutral lines in the key).

A1 and A2 titration curves

Dendrimer A1 and A2 were re-synthesized following the procedure described in reference 47 and their titration curves determined using the procedure described in the previous reference. Briefly, the dendrimers were titrated by adding aliquots 20 µL of 2 mM NaOH into a 1 mM aqueous solution of dendrimer, at 25 °C. The experiments were repeated and gave reproducible values within 10% error. The mass spectra (ESI+) for A1 and A2 are presented in Figures D.11 and D.12.

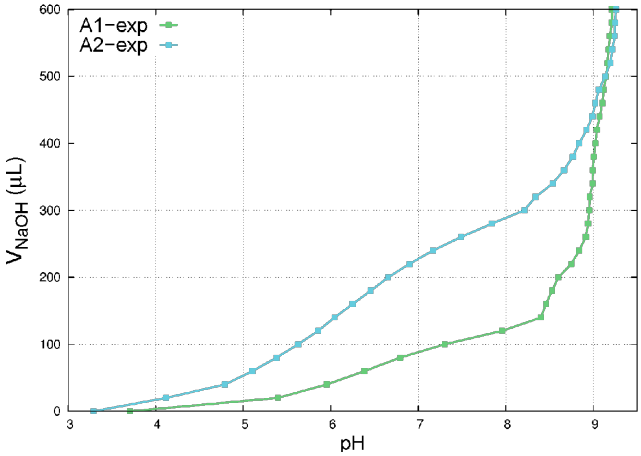


Figure D.10: Titration curves for A1 and A2. A dendrimer solution (1 mM) was titrated with 20 µL aliquots of a NaOH solution (2 mM).

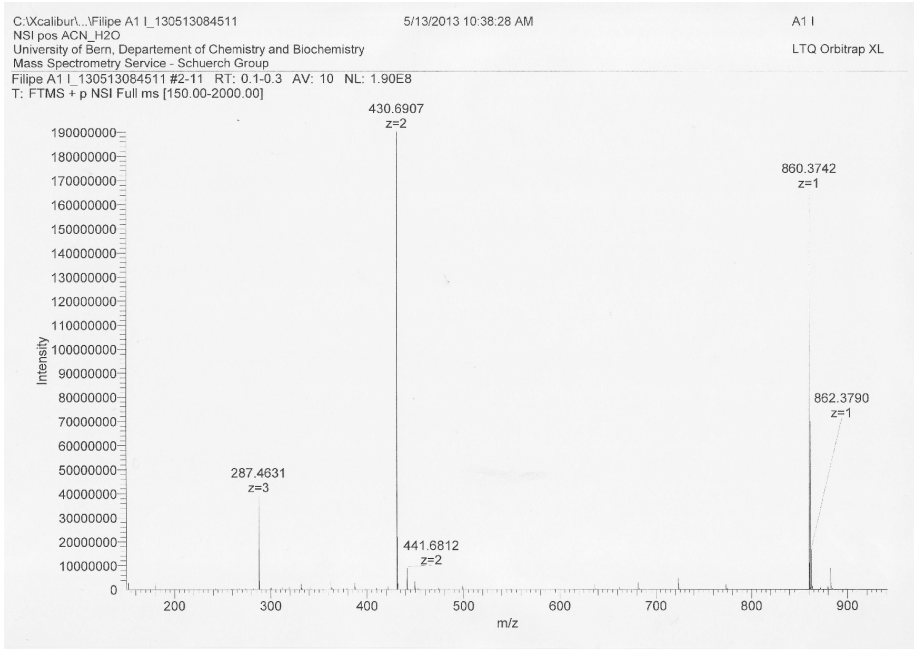


Figure D.11: Mass Spectra of A1. MS (ESI+) calc. for $C_{34}H_{49}N_{15}O_{12}$ $[M+H]^+$: 859.37, found: 860.37.

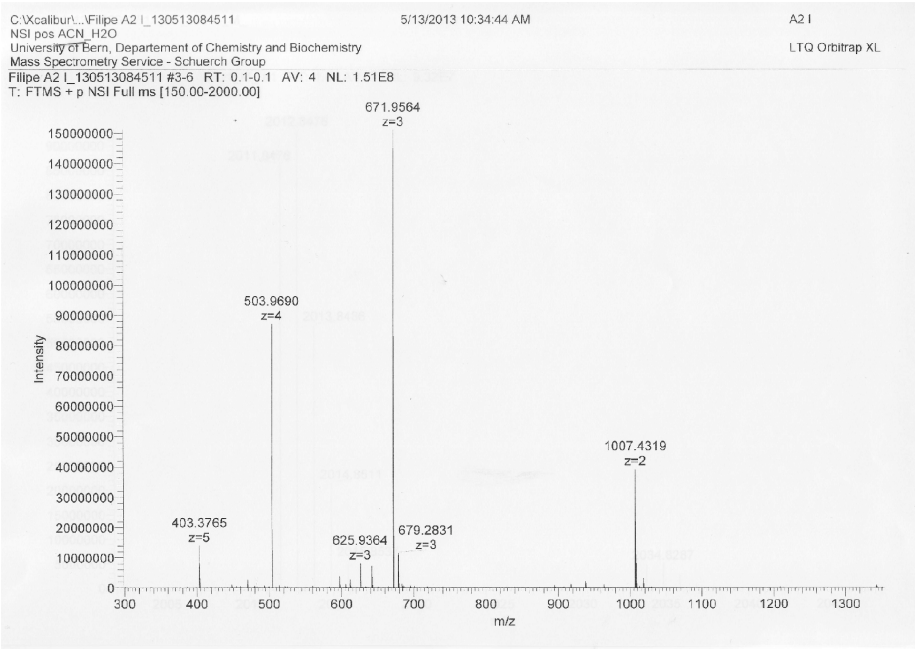


Figure D.12: Mass Spectra of A2. MS (ESI+) calc. for $C_{80}H_{113}N_{35}O_{28}$ $[M+H]^+$: 2011.84, found: 2012.85.

R_g histograms at different pH values

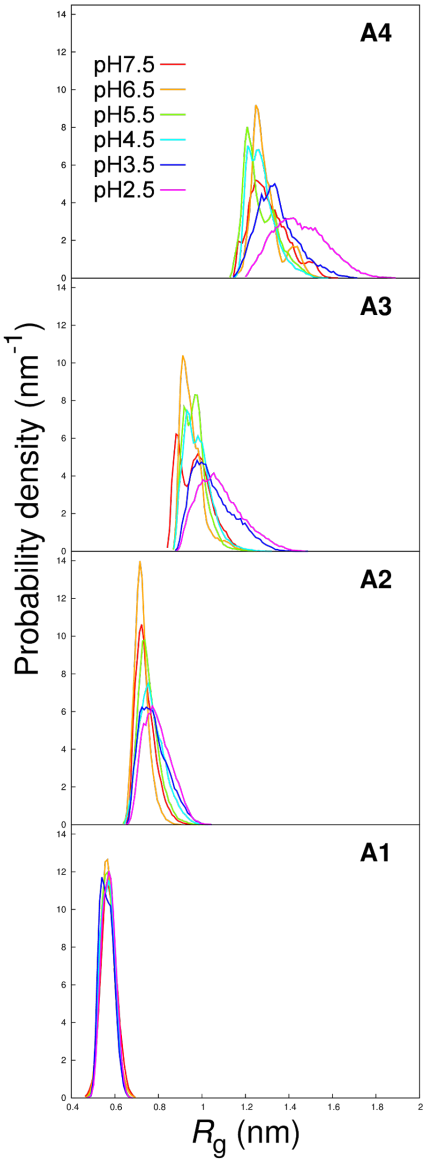


Figure D.13: R_g probability density histograms for A1, A2, A3 and A4 at different pH values.

Average SASA at different pH values

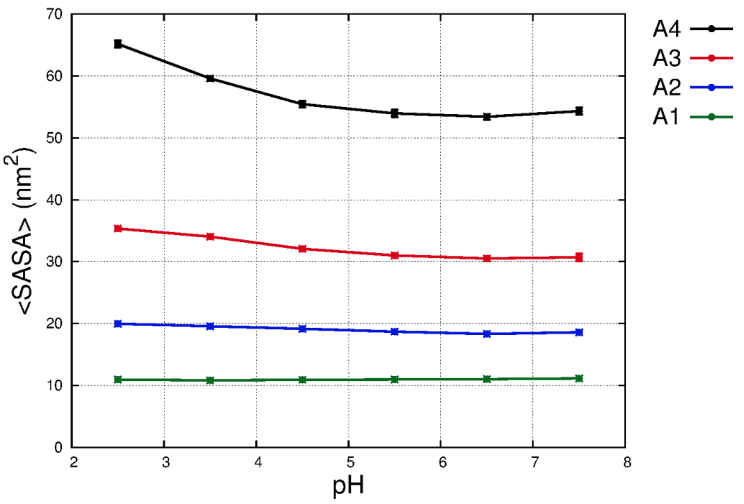


Figure D.14: Average solvent accessible surface area of the dendrimer at different pH values.

Energy landscapes: A1, A2 and A3

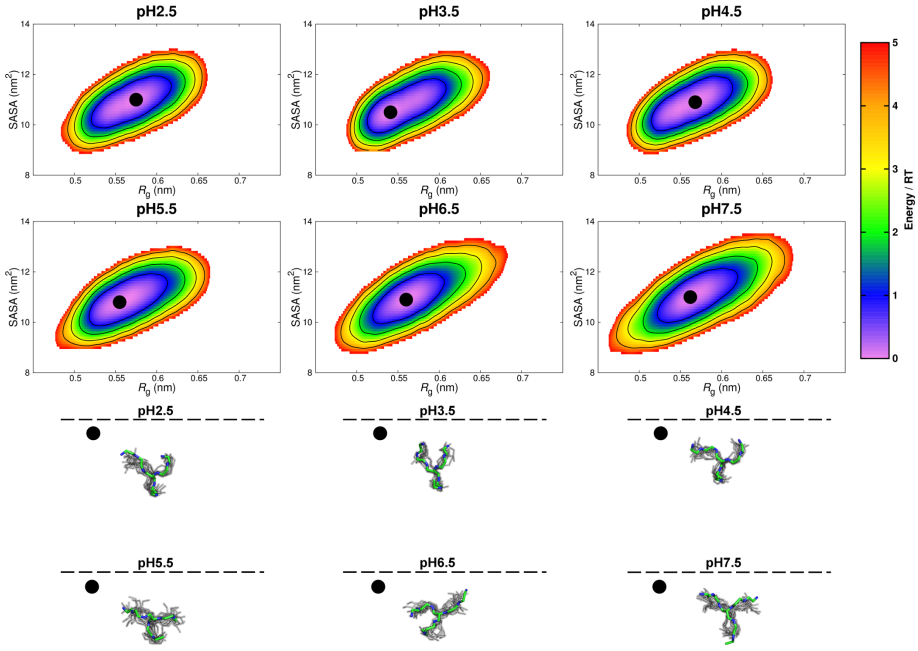


Figure D.15: Dendrimer A1 energy profiles and lowest energy conformers at different pH values. The energy profiles use R_g and SASA as structural coordinates. In the energy profiles the lowest energy cluster is represented as a black dot. For the different pH values pictures of the 21 lowest energy conformers of each energy cluster are shown, with the lowest energy structure from each cluster highlighted (only the peptidic backbone is represented using green for carbons and blue for nitrogen atoms) and superimposed on the remaining 20 structures (in gray).

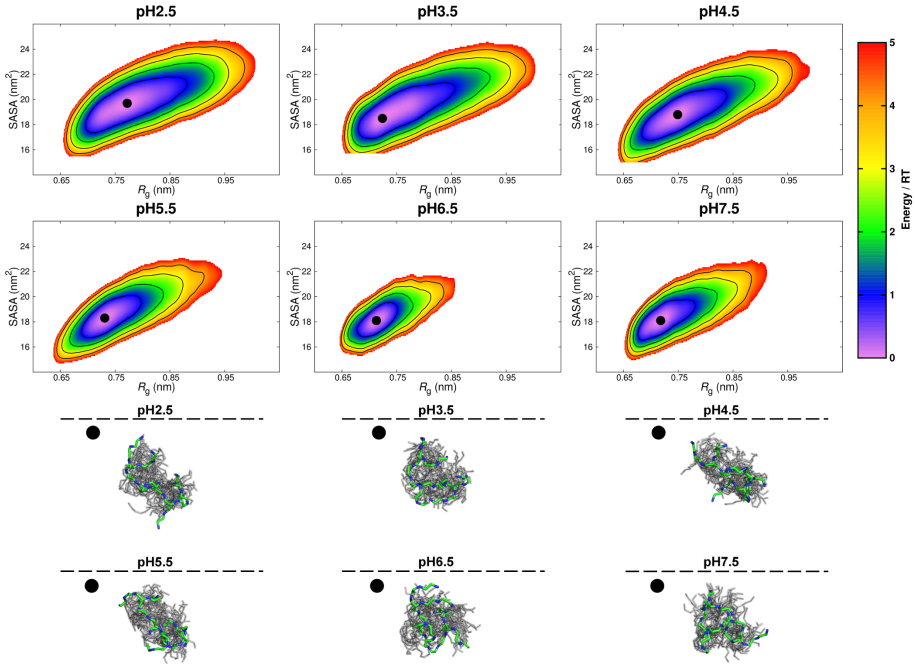


Figure D.16: Dendrimer A2 energy profiles and lowest energy conformers at different pH values. The energy profiles use R_g and SASA as structural coordinates. In the energy profiles the lowest energy cluster is represented as a black dot. For the different pH values pictures of the 21 lowest energy conformers of each energy cluster are shown, with the lowest energy structure from each cluster highlighted (only the peptidic backbone is represented using green for carbons and blue for nitrogen atoms) and superimposed on the remaining 20 structures (in gray).

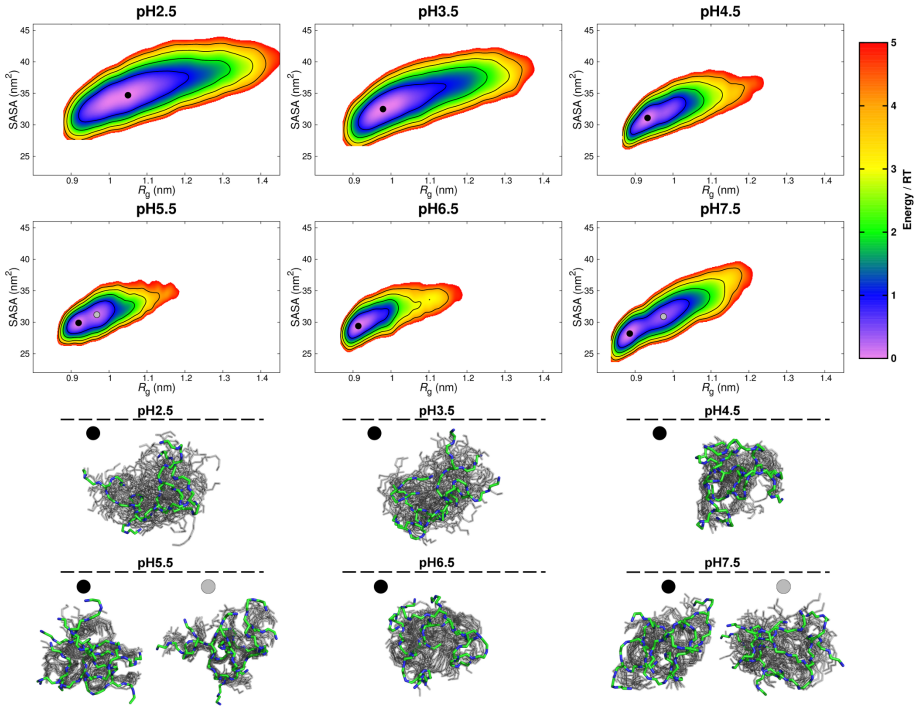


Figure D.17: Dendrimer A3 energy profiles and lowest energy conformers at different pH values. The energy profiles use R_g and SASA as structural coordinates. In the energy profiles the lowest energy cluster is represented as a black dot. For the different pH values pictures of the 21 lowest energy conformers of each energy cluster are shown, with the lowest energy structure from each cluster highlighted (only the peptidic backbone is represented using green for carbons and blue for nitrogen atoms) and superimposed on the remaining 20 structures (in gray).

π-π interactions

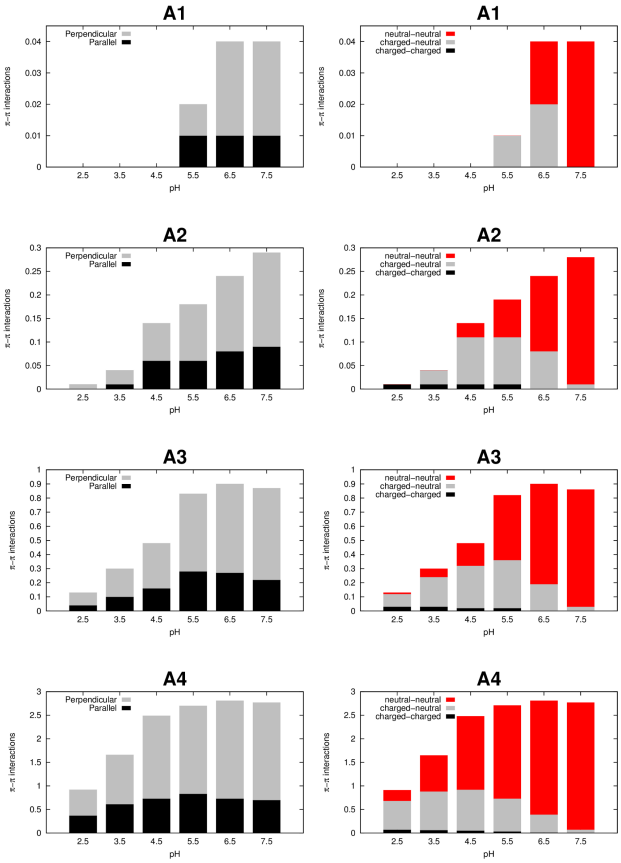


Figure D.18: Distribution of π-π stacking by type (left) and histidine charged states (right). The sum the values in each column yields the average number of π-π interactions at each pH.

Hydrogen Bonds: MainChain vs. SideChain

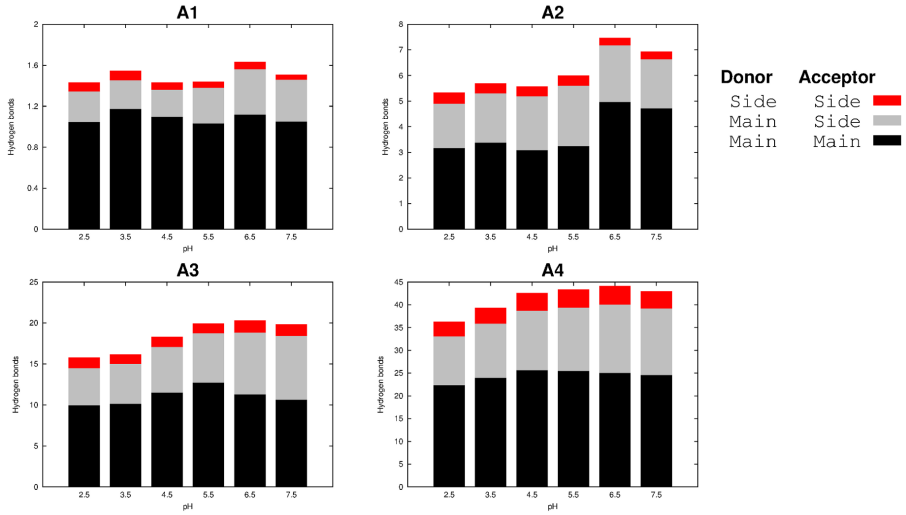


Figure D.19: Distribution of HB by donors and acceptors of the dendrimers main-chain or side-chain. The sum the values in each column yields the average number of HB at each pH.

Histidines protonation curves

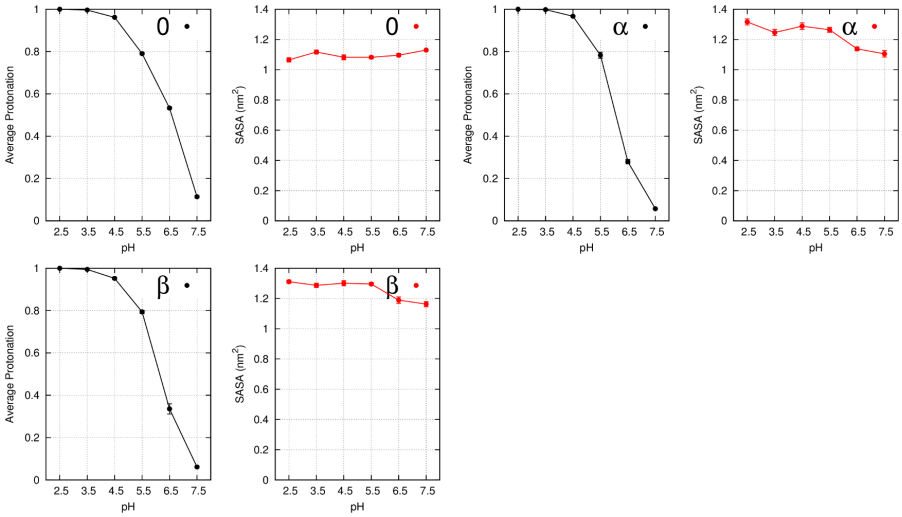


Figure D.20: Dendrimer A1. Histidines average proton occupancy and imidazole ring solvent accessible surface area curves.

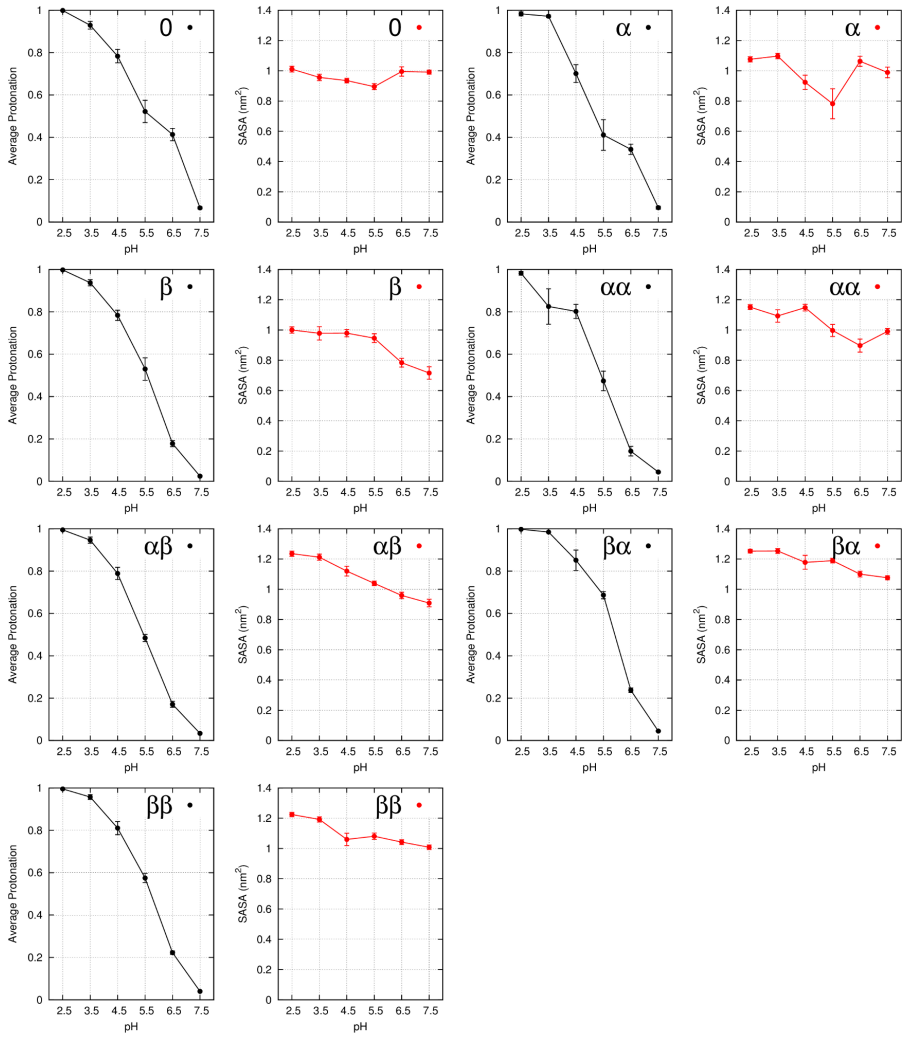
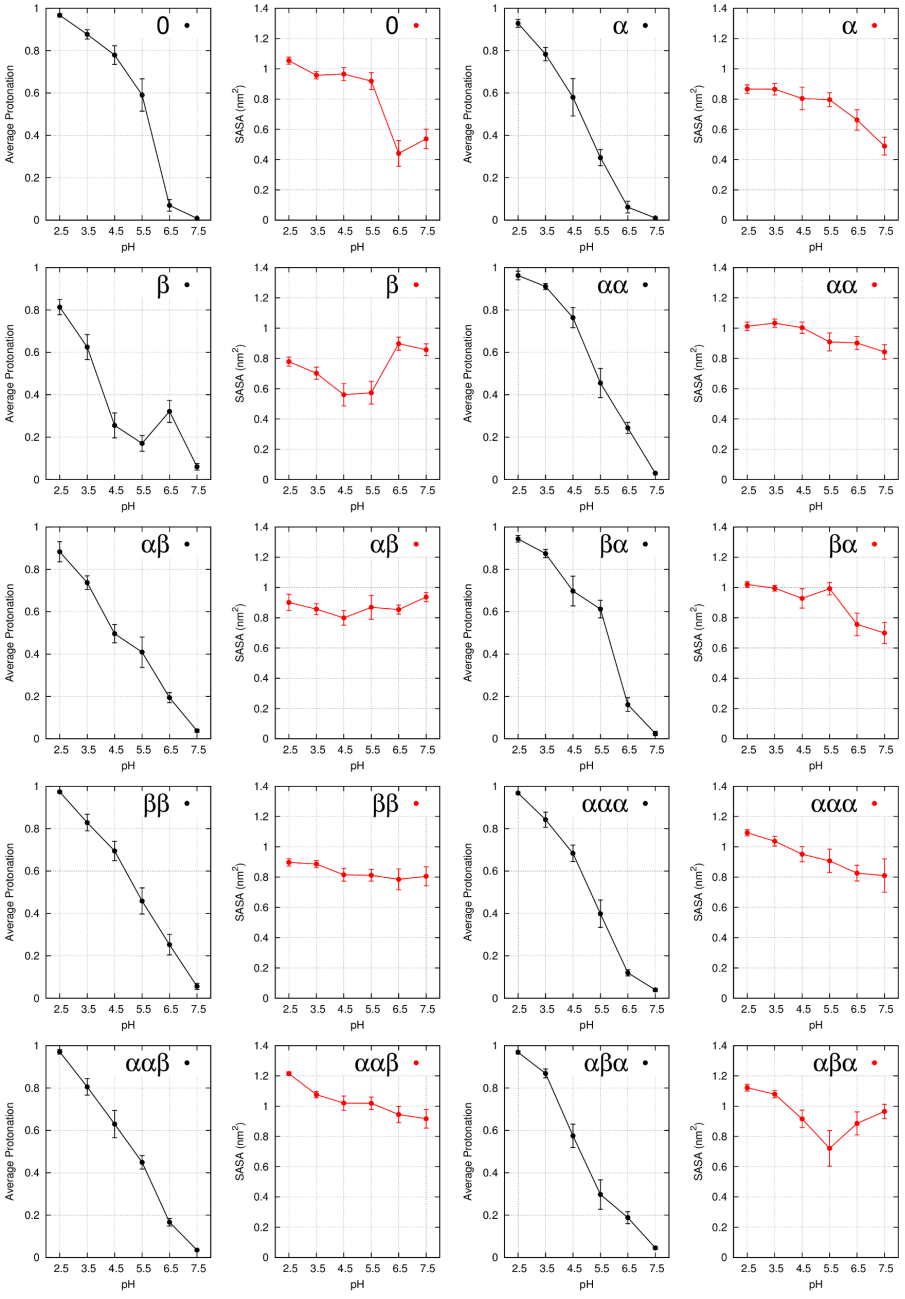


Figure D.21: Dendrimer A2. Histidines average proton occupancy and imidazole ring solvent accessible surface area curves.


0
0
α
α
β
β
αα
αα
αβ
αβ
βα
βα
ββ
ββ
ααα
ααα
ααβ
ααβ
αβα
αβα
Average Protonation
SASA (nm²)
pH
2.5 3.5 4.5 5.5 6.5 7.5

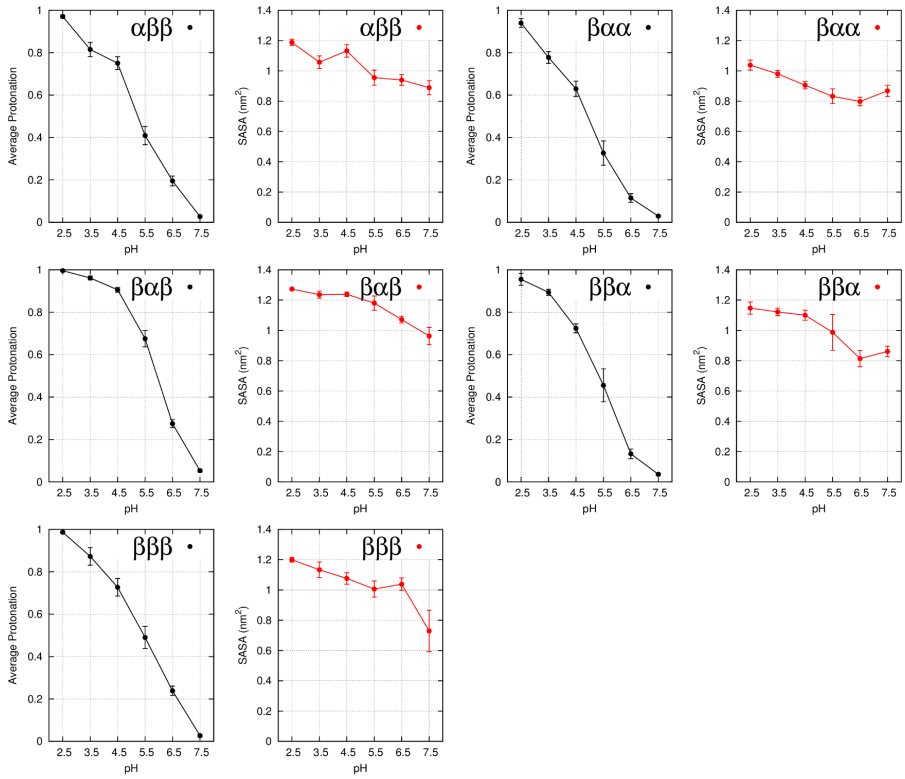
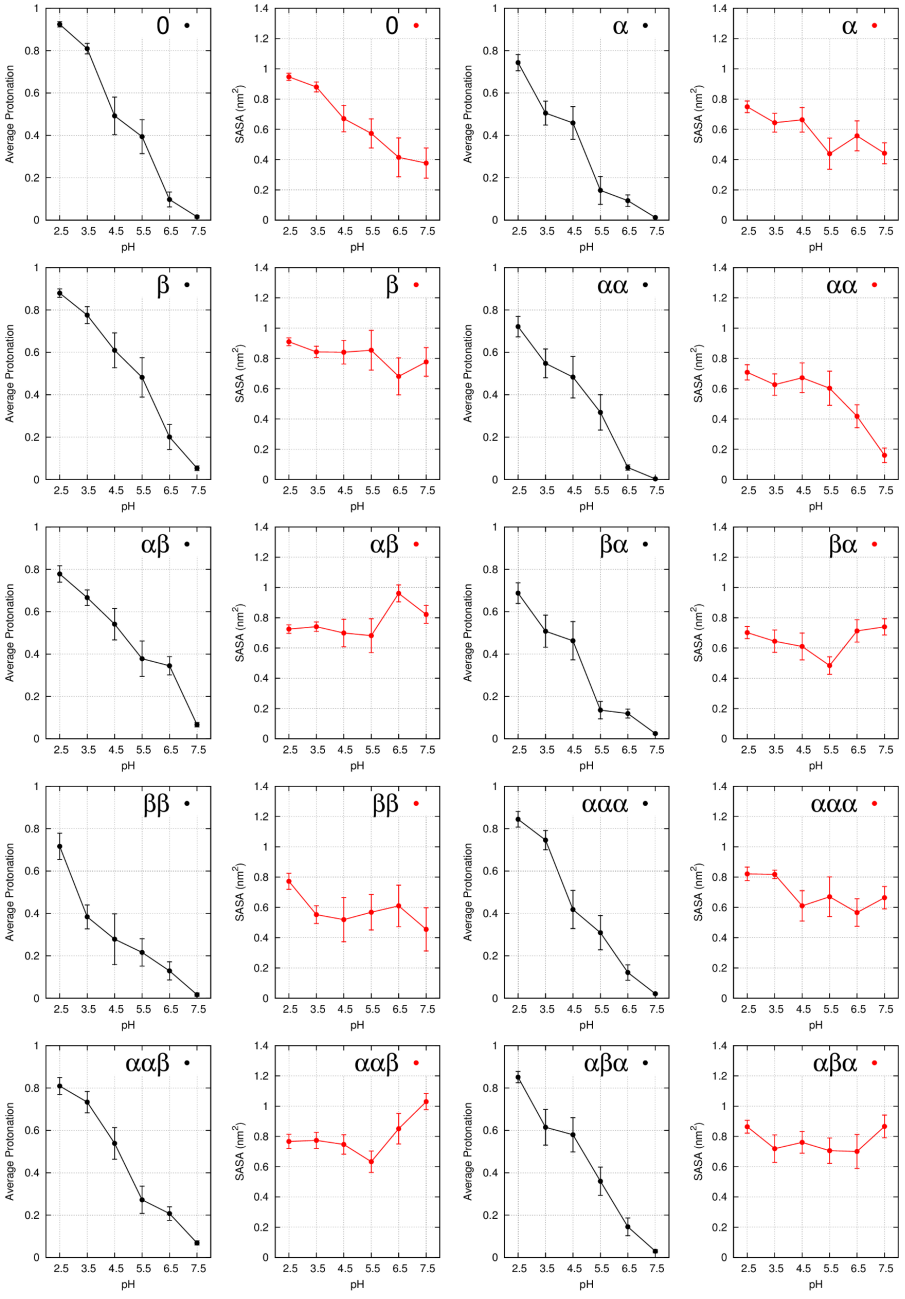
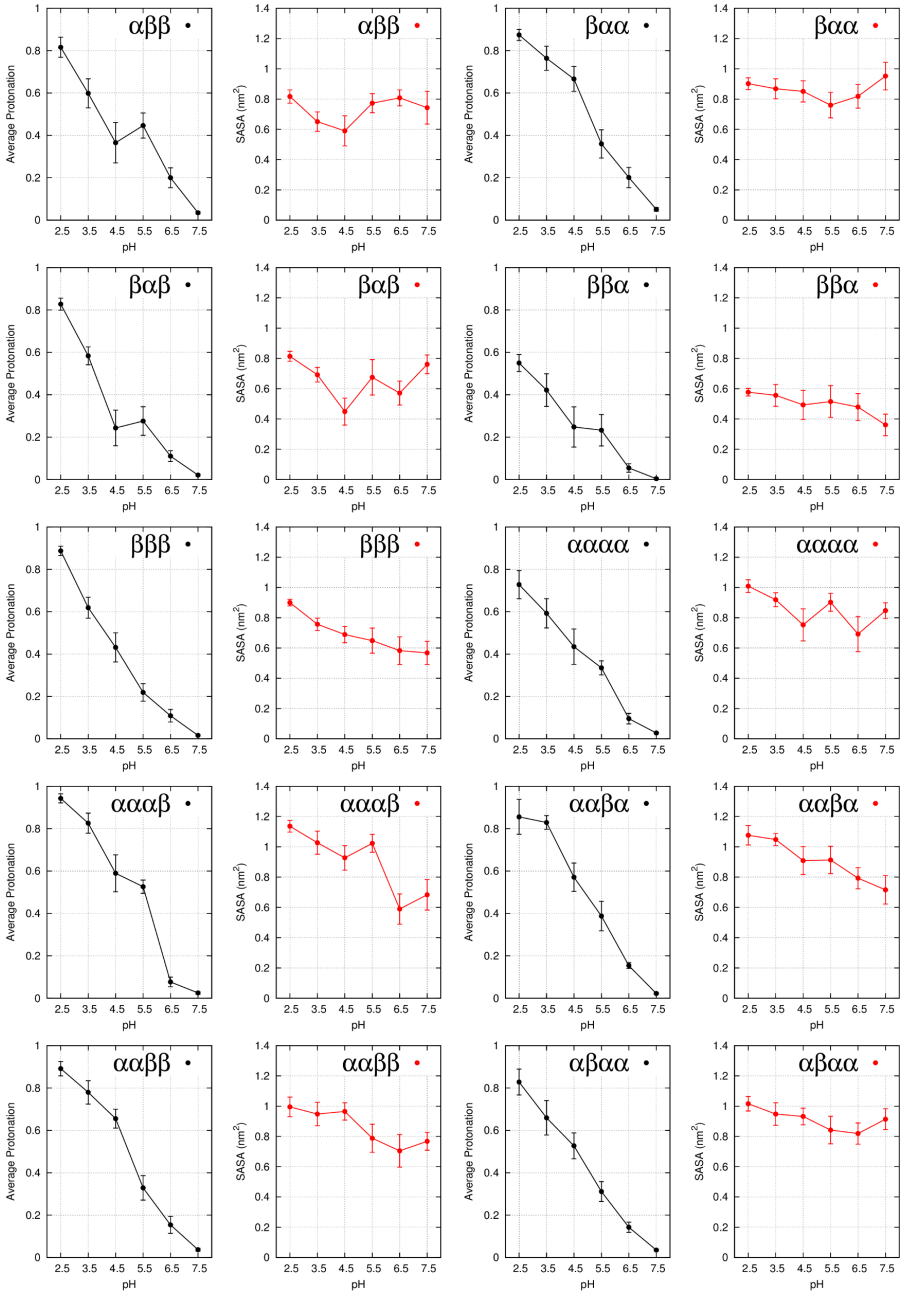
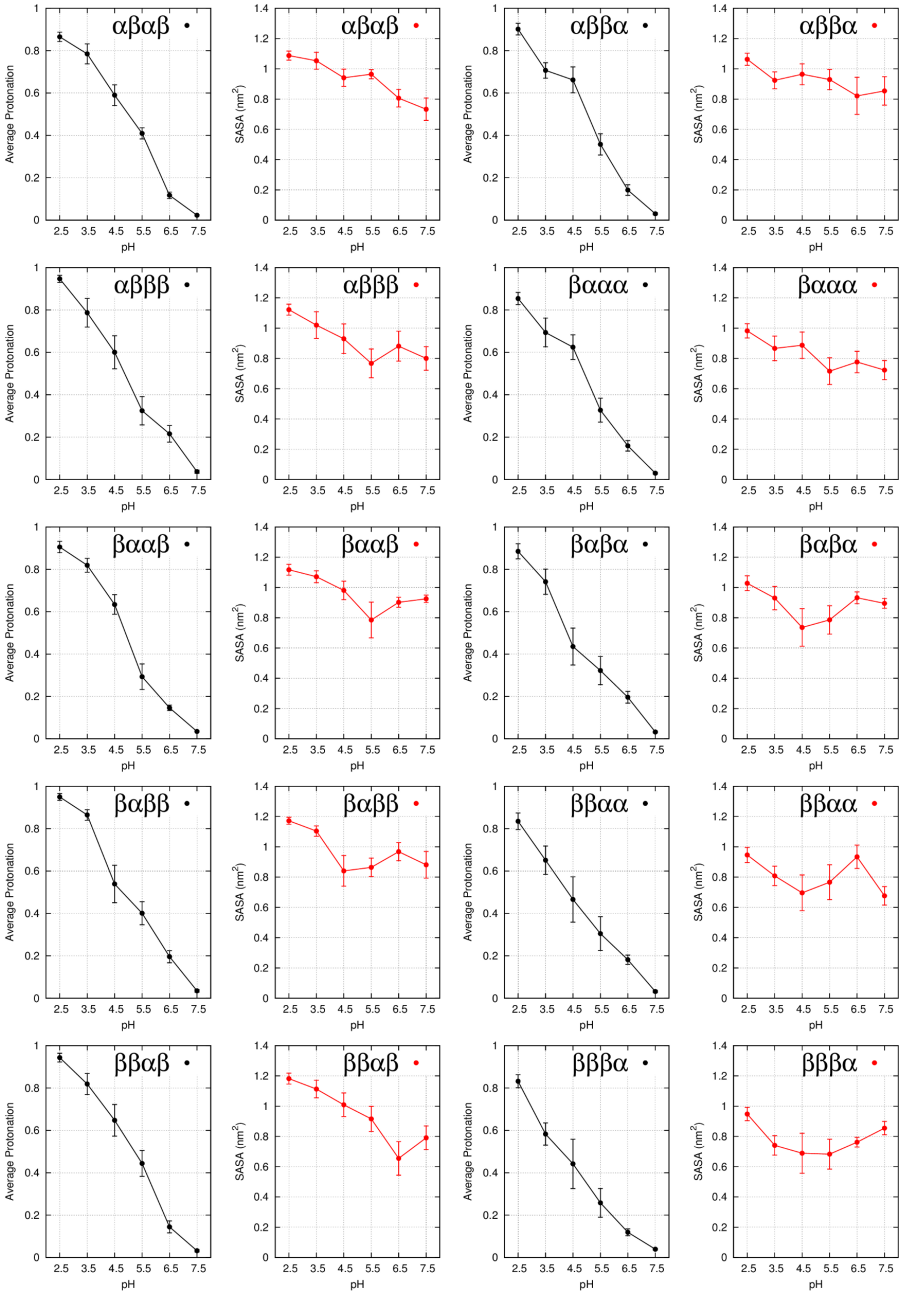


Figure D.22: Dendrimer A3. Histidines average proton occupancy and imidazole ring solvent accessible surface area curves.


0
0
α
α
β
β
αα
αα
αβ
αβ
βα
βα
ββ
ββ
ααα
ααα
ααβ
ααβ
αβα
αβα
Average Protonation
SASA (nm²)
pH


αββ
αββ
βαα
βαα
βαβ
βαβ
ββα
ββα
βββ
βββ
αααα
αααα
αααβ
αααβ
ααβα
ααβα
αάββ
αάββ
αβαα
αβαα
Average Protonation
SASA (nm2)
pH
2.5 3.5 4.5 5.5 6.5 7.5


αβαβ
αβαβ
αββα
αββα
αβββ
αβββ
βααα
βααα
ββαβ
ββαβ
βαβα
βαβα
βαββ
βαββ
ββαα
ββαα
ββαβ
ββαβ
βββα
βββα
Average Protonation
SASA (nm²)
pH
2.5 3.5 4.5 5.5 6.5 7.5

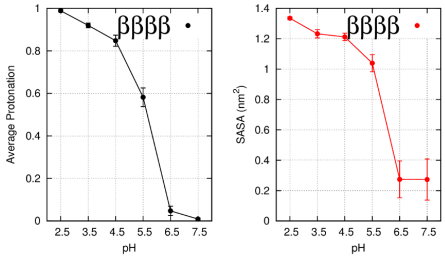


Figure D.23: Dendrimer A4. Histidines average proton occupancy and imidazole ring solvent accessible surface area curves.

Dendrimer-Substrate contact area

The percentage of contact area between a dendrimer and a molecule of substrate was calculated as:

$$ContactArea = \left(1 - \frac{\text{SASA}_{\text{complex}}}{\text{SASA}_{\text{free}}}\right) \times 100 \tag{D.1}$$

where, $\text{SASA}_{\text{complex}}$ is the solvent accessible surface area of the substrate when considering the dendrimer-substrate complex and $\text{SASA}_{\text{free}}$ is the total solvent accessible surface area of the substrate alone. Hence, the *ContactArea* is the percentage of substrate SASA that is in contact with the dendrimer surface. This value was computed for each substrate in each frame. SASA values were obtained using the method described in reference 242 using a 1.4 Å spherical probe.

Dendrimer-substrate interaction tests

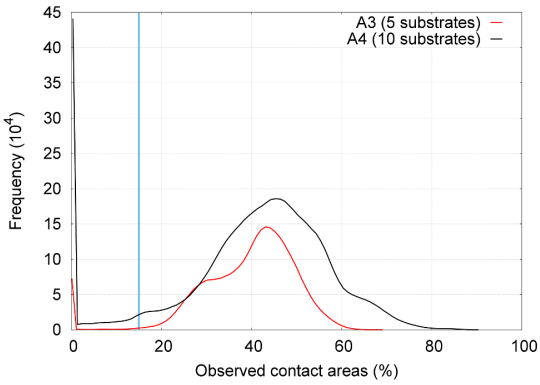


Figure D.24: Distribution of dendrimer–substrate contact areas. The selected cutoff (15% of contact area) is displayed as a blue line.

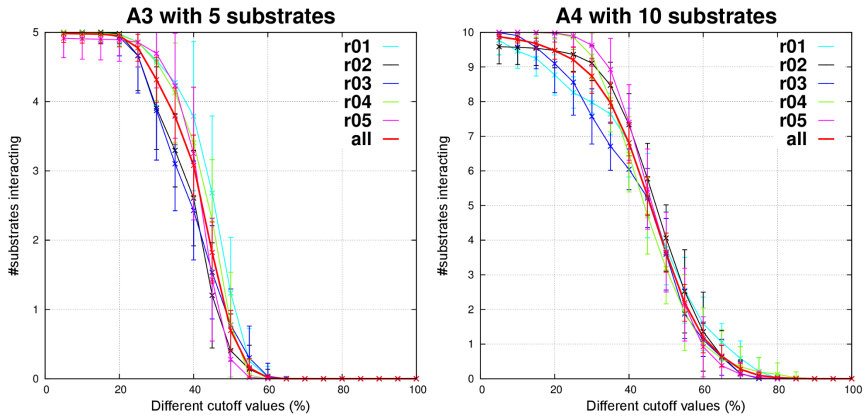


Figure D.25: Number of substrate molecules considered as interacting with the dendrimer when testing different cutoff values for the contact area. Curves for each replicate and for the entire concatenated trajectories (all) are presented.

Dendrimer-Substrate complexes

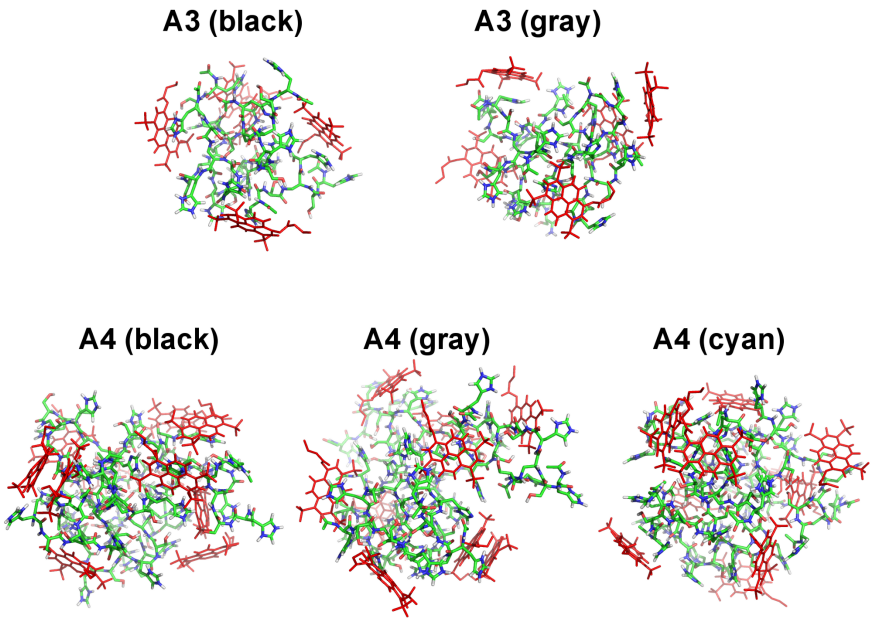


Figure D.26: Examples of dendrimer-substrate complexes. The images correspond to the lowest energy conformations of the 2D landscapes of A3 and A4 in the presence of 5 and 10 substrate molecules, respectively (Figure 9 in the article). Substrate molecules are shown in red.

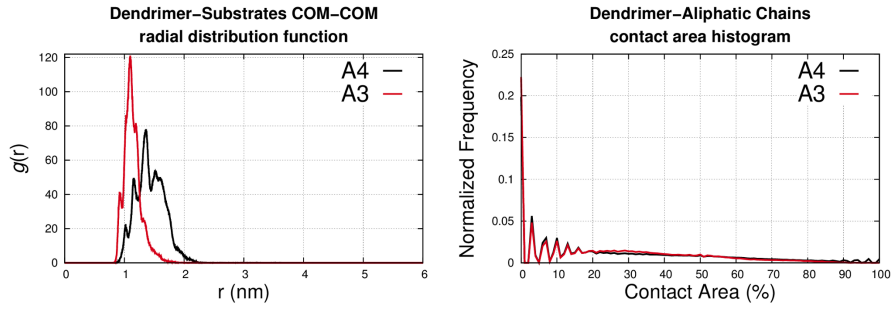


Figure D.27: (Left) Dendrimer–substrates COM-COM radial distribution functions. (Right) Histogram of the contact area between the aliphatic chains of the substrate molecules and the dendrimer; binsize = 1%.

Bibliography

[1] F. Vögtle, G. Richardt, N. Werner, and A. J. Rackstraw, (2009), *Dendrimer Chemistry*. Wiley.

[2] J. M. J. Fréchet and D. A. Tomalia, (2001), *Dendrimers and other dendritic polymers*. Wiley series in polymer science, Wiley.

[3] D. A. Tomalia, (2005), "The dendritic state", *Mater. Today*, **8**, 34–46.

[4] D. A. Tomalia, A. M. Naylor, and W. A. Goddard, (1990), "Starburst dendrimers: molecular-level control of size, shape, surface chemistry, topology, and flexibility from atoms to macroscopic matter", *Angew. Chem., Int. Ed.*, **29**, 138–175.

[5] R. Esfand and D. A. Tomalia, (2001), "Poly (amidoamine)(PAMAM) dendrimers: from biomimicry to drug delivery and biomedical applications", *Drug Discovery Today*, **6**, 427–436.

[6] S. Campagna, P. Ceroni, and F. Puntoriero, (2011), *Designing Dendrimers*. Wiley.

[7] S. M. Grayson and J. M. J. Fréchet, (2001), "Convergent dendrons and dendrimers: from synthesis to applications", *Chem. Rev.*, **101**, 3819–3868.

[8] B. Helms and E. Meijer, (2006), "Dendrimers at work", *Science*, **313**, 929–930.

[9] C. C. Lee, J. A. MacKay, J. M. J. Frechet, and F. C. Szoka, (2005), "Designing dendrimers for biological applications", *Nat. Biotech.*, **23**, 1517–1526.

[10] S. H. Medina and M. E. H. El-Sayed, (2009), "Dendrimers as carriers for delivery of chemotherapeutic agents", *Chem. Rev.*, **109**, 3141–3157.

[11] P. A. Gunatillake, G. Odian, and D. A. Tomalia, (1988), "Thermal polymerization of a 2-(carboxyalkyl)-2-oxazoline", *Macromolecules*, **21**, 1556–1562.

[12] M. Gauthier and M. Moeller, (1991), "Uniform highly branched polymers by anionic grafting: arborescent graft polymers", *Macromolecules*, **24**, 4548–4553.

[13] E. R. Gillies and J. M. J. Fréchet, (2005), "Dendrimers and dendritic polymers in drug delivery", *Drug Discovery Today*, **10**, 35–43.

[14] G. R. Newkome, C. N. Moorefield, and F. Vögtle, (2001), *Dendrimers and Dendrons: Concepts, Syntheses, Applications*. Wiley-VCH.

[15] M. Liu and J. J. M. J. Fréchet, (1999), "Designing dendrimers for drug delivery", *Pharm. Sci. Technol. Today*, **2**, 393–401.

[16] A. W. Bosman, H. M. Janssen, and E. W. Meijer, (1999), "About dendrimers: Structure, physical properties, and applications", *Chem. Rev.*, **99**, 1665–1688.

[17] U. Boas and P. M. Heegaard, (2004), "Dendrimers in drug research", *Chem. Soc. Rev.*, **33**, 43–63.

[18] A.-M. Caminade, S. Fruchon, C.-O. Turrin, M. Poupot, A. Ouali, A. Maraval, M. Garzoni, M. Maly, V. Furer, V. Kovalenko, *et al.*, (2015), "The key role of the scaffold on the efficiency of dendrimer nanodrugs", *Nat. Commun.*, **6**, 7722.

[19] D. A. Tomalia, (2012), "Dendritic effects: dependency of dendritic nano-periodic property patterns on critical nanoscale design parameters (CNDPs)", *New J. Chem.*, **36**, 264–281.

[20] G. M. Dykes, (2001), "Dendrimers: a review of their appeal and applications", *J. Chem. Technol. Biotechnol.*, **76**, 903–918.

[21] A. R. Menjoge, R. M. Kannan, and D. A. Tomalia, (2010), "Dendrimer-based drug and imaging conjugates: design considerations for nanomedical applications", *Drug Discovery Today*, **15**, 171–185.

[22] D. K. Smith and F. Diederich, (1998), "Functional dendrimers: unique biological mimics", *Chem. - Eur. J.*, **4**, 1353–1361.

[23] D. Astruc, E. Boisselier, and C. Ornelas, (2010), "Dendrimers designed for functions: From physical, photophysical, and supramolecular properties to applications in sensing, catalysis, molecular electronics, photonics, and nanomedicine", *Chem. Rev.*, **110**, 1857–1959.

[24] Y. Cheng, L. Zhao, Y. Li, and T. Xu, (2011), "Design of biocompatible dendrimers for cancer diagnosis and therapy: current status and future perspectives", *Chem. Soc. Rev.*, **40**, 2673–2703.

[25] S. Svenson and D. A. Tomalia, (2005), "Dendrimers in biomedical applications–reflections on the field", *Adv. Drug Delivery Rev.*, **57**, 2106–2129.

[26] P. J. Flory, (1941), "Molecular size distribution in three dimensional polymers. I. gelation", *J. Am. Chem. Soc.*, **63**, 3083–3090.

[27] P. J. Flory, (1941), "Molecular size distribution in three dimensional polymers. II. trifunctional branching units", *J. Am. Chem. Soc.*, **63**, 3091–3096.

[28] P. J. Flory, (1941), "Molecular size distribution in three dimensional polymers. III. tetrafunctional branching units", *J. Am. Chem. Soc.*, **63**, 3096–3100.

[29] P. J. Flory, (1952), "Molecular size distribution in three dimensional polymers. VI. branched polymers containing A-R-B_{f-1} type units", *J. Am. Chem. Soc.*, **74**, 2718–2723.

[30] W. H. Stockmayer, (1943), "Theory of molecular size distribution and gel formation in branched-chain polymers", *J. Chem. Phys.*, **11**, 45–55.

[31] E. Buhleier, W. Wehner, and F. Vögtle, (1978), ""Cascade"-and" Nonskid-Chain-like" syntheses of molecular cavity topologies", *Synthesis*, **1978**, 155–158.

[32] R. G. Denkewalter, J. Kolc, and W. J. Lukasavage, (1981), "Macromolecular highly branched homogeneous compound based on lysine units", US Patent 4,289,872.

[33] D. A. Tomalia, H. Baker, J. Dewald, M. Hall, G. Kallos, S. Martin, J. Roeck, J. Ryder, and P. Smith, (1985), "A new class of polymers: starburst-dendritic macromolecules", *Polym. J.*, **17**, 117–132.

[34] C. J. Hawker and J. M. J. Frechet, (1990), "Preparation of polymers with controlled molecular architecture. a new convergent approach to dendritic macromolecules", *J. Am. Chem. Soc.*, **112**, 7638–7647.

[35] T. Kimmerlin and D. Seebach, (2005), ""100 years of peptide synthesis": ligation methods for peptide and protein synthesis with applications to β-peptide assemblies", *J. Pept. Res.*, **65**, 229–260.

[36] K. Sadler and J. P. Tam, (2002), "Peptide dendrimers: applications and synthesis", *Rev. Mol. Biotechnol.*, **90**, 195–229.

[37] L. Crespo, G. Sanclimens, M. Pons, E. Giralt, M. Royo, and F. Albericio, (2005), "Peptide and amide bond-containing dendrimers", *Chem. Rev.*, **105**, 1663–1682.

[38] J. F. G. A. Jansen, E. W. Brabander, Ellen M. M. Meijer, *et al.*, (1994), "Encapsulation of guest molecules into a dendritic box", *Science*, **266**, 1226–1229.

[39] E. Oledzka, P. Sliwerska, M. Sobczak, B. Kraska, W. Kamysz, G. Nalecz-Jawecki, and W. Kolodziejski, (2015), "Peptide dendrimer functionalized with amphiphilic triblock copolymers: Synthesis and characterization", *Macromol. Chem. Phys.*, **216**, 1365–1375.

[40] P. Niederhafner, J. Šebestík, and J. Ježek, (2005), "Peptide dendrimers", *J. Pept. Sci.*, **11**, 757–788.

[41] T. Darbre and J.-L. Reymond, (2006), "Peptide dendrimers as artificial enzymes, receptors, and drug-delivery agents", *Acc. Chem. Res.*, **39**, 925–934.

[42] J.-L. Reymond and T. Darbre, (2012), "Peptide and glycopeptide dendrimer apple trees as enzyme models and for biomedical applications", *Org. Biomol. Chem.*, **10**, 1483–1492.

[43] W. B. Kauffman, T. Fuselier, J. He, and W. C. Wimley, (2015), "Mechanism matters: A taxonomy of cell penetrating peptides", *Trends Biochem. Sci.*, **40**, 749–764.

[44] A. Kwok, G. A. Eggimann, J.-L. Reymond, T. Darbre, and F. Hollfelder, (2013), "Peptide dendrimer/lipid hybrid systems are efficient DNA transfection reagents: Structure-activity relationships highlight the role of charge distribution across dendrimer generations", *ACS Nano*, **7**, 4668–4682.

[45] A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K.-E. Jaeger, M. Lahmann, *et al.*, (2013), "Multivalent glycoconjugates as anti-pathogenic agents", *Chem. Soc. Rev.*, **42**, 4709–4727.

[46] E. Delort, T. Darbre, and J.-L. Reymond, (2004), "A strong positive dendritic effect in a peptide dendrimer-catalyzed ester hydrolysis reaction", *J. Am. Chem. Soc.*, **126**, 15642–15643.

[47] E. Delort, N.-Q. Nguyen-Trung, T. Darbre, and J.-L. Reymond, (2006), "Synthesis and activity of histidine-containing catalytic peptide dendrimers", *J. Org. Chem.*, **71**, 4468–4480.

[48] P. Sommer, V. S. Fluxa, T. Darbre, and J.-L. Reymond, (2009), "Proteolysis of peptide dendrimers", *ChemBioChem*, **10**, 1527–1536.

[49] S. B. Kent, (1988), "Chemical synthesis of peptides and proteins", *Annu. Rev. Biochem.*, **57**, 957–989.

[50] N. A. Uhlich, T. Darbre, and J.-L. Reymond, (2011), "Peptide dendrimer enzyme models for ester hydrolysis and aldolization prepared by convergent thioether ligation", *Org. Biomol. Chem.*, **9**, 7071–7084.

[51] K. S. Lam, M. Lebl, and V. Krchnák, (1997), "The "one-bead-one-compound" combinatorial library method", *Chem. Rev.*, **97**, 411–448.

[52] N. Maillard, A. Clouet, T. Darbre, and J.-L. Reymond, (2009), "Combinatorial libraries of peptide dendrimers: design, synthesis, on-bead high-throughput screening, bead decoding and characterization", *Nat. Protoc.*, **4**, 132–142.

[53] A. Clouet, T. Darbre, and J.-L. Reymond, (2004), "A combinatorial approach to catalytic peptide dendrimers", *Angew. Chem.*, **116**, 4712–4715.

[54] J. Kofoed, T. Darbre, and J.-L. Reymond, (2006), "Artificial aldolases from peptide dendrimer combinatorial libraries", *Org. Biomol. Chem.*, **4**, 3268–3281.

[55] P. Sommer, N. A. Uhlich, J.-L. Reymond, and T. Darbre, (2008), "A peptide dendrimer model for vitamin B12 transport proteins", *ChemBioChem*, **9**, 689–693.

[56] J.-L. Reymond, M. Bergmann, and T. Darbre, (2013), "Glycopeptide dendrimers as *Pseudomonas aeruginosa* biofilm inhibitors", *Chem. Soc. Rev.*, **42**, 4814–4822.

[57] A. Esposito, E. Delort, D. Lagnoux, F. Djojo, and J.-L. Reymond, (2003), "Catalytic peptide dendrimers", *Angew. Chem., Int. Ed.*, **42**, 1381–1383.

[58] C. Douat-Casassus, T. Darbre, and J.-L. Reymond, (2004), "Selective catalysis with peptide dendrimers", *J. Am. Chem. Soc.*, **126**, 7817–7826.

[59] S. Javor, E. Delort, T. Darbre, and J.-L. Reymond, (2007), "A peptide dendrimer enzyme model with a single catalytic site at the core", *J. Am. Chem. Soc.*, **129**, 13238–13246.

[60] N. Maillard, R. Biswas, T. Darbre, and J.-L. Reymond, (2011), "Combinatorial discovery of peptide dendrimer enzyme models hydrolyzing isobutyryl fluorescein", *ACS Comb. Sci.*, **13**, 310–320.

[61] J. Kofoed and J.-L. Reymond, (2007), "A general method for designing combinatorial peptide libraries decodable by amino acid analysis", *J. Comb. Chem.*, **9**, 1046–1052.

[62] D. Lagnoux, T. Darbre, M. L. Schmitz, and J.-L. Reymond, (2005), "Inhibition of mitosis by glycopeptide dendrimer conjugates of colchicine", *Chem. - Eur. J.*, **11**, 3941–3950.

[63] E. M. V. Johansson, J. Dubois, T. Darbre, and J.-L. Reymond, (2010), "Glycopeptide dendrimer colchicine conjugates targeting cancer cells", *Bioorg. Med. Chem.*, **18**, 6589–6597.

[64] N. A. Uhlich, A. Natalello, R. U. Kadam, S. M. Doglia, J.-L. Reymond, and T. Darbre, (2010), "Structure and binding of peptide-dendrimer ligands to vitamin B12", *ChemBioChem*, **11**, 358–365.

[65] N. A. Uhlich, P. Sommer, C. Buhr, S. Schurch, J.-L. Reymond, and T. Darbre, (2009), "Remote control of bipyridine-metal coordination within a peptide dendrimer", *Chem. Commun.*, 6237–6239.

[66] P. Geotti-Bianchini, T. Darbre, and J.-L. Reymond, (2013), "pH-tuned metal coordination and peroxidase activity of a peptide dendrimer enzyme model with a Fe(II) bipyridine at its core", *Org. Biomol. Chem.*, **11**, 344–352.

[67] E. M. V. Johansson, S. A. Crusz, E. Kolomiets, L. Buts, R. U. Kadam, M. Cacciarini, K.-M. Bartels, S. P. Diggle, M. Cámara, P. Williams, *et al.*, (2008), "Inhibition and dispersion of *Pseudomonas aeruginosa* biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB", *Chem. Biol. (Oxford, U. K.)*, **15**, 1249 – 1257.

[68] E. Kolomiets, M. A. Swiderska, R. U. Kadam, E. Johansson, K.-E. Jaeger, T. Darbre, and J.-L. Reymond, (2009), "Glycopeptide dendrimers with high affinity for the fucose-binding lectin LecB from *Pseudomonas aeruginosa*", *ChemMedChem*, **4**, 562–569.

[69] E. M. V. Johansson, R. U. Kadam, G. Rispoli, S. A. Crusz, K.-M. Bartels, S. P. Diggle, M. Camara, P. Williams, K.-E. Jaeger, T. Darbre, and J.-L. Reymond, (2011), "Inhibition of *Pseudomonas aeruginosa* biofilms with a glycopeptide dendrimer containing D-amino acids", *Med. Chem. Commun.*, **2**, 418–420.

[70] R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, *et al.*, (2011), "A glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of *Pseudomonas aeruginosa* biofilms", *Angew. Chem., Int. Ed.*, **50**, 10631–10635.

[71] M. Bergmann, G. Michaud, R. Visini, X. Jin, E. Gillon, A. Stocker, A. Imberty, T. Darbre, and J.-L. Reymond, (2016), "Multivalency effects on *Pseudomonas aeruginosa* biofilm inhibition and dispersal by glycopeptide dendrimers targeting lectin leca", *Org. Biomol. Chem.*, **14**, 138–148.

[72] M. Stach, N. Maillard, R. U. Kadam, D. Kalbermatter, M. Meury, M. G. P. Page, D. Fotiadis, T. Darbre, and J.-L. Reymond, (2012), "Membrane disrupting antimicrobial peptide dendrimers with multiple amino termini", *Med. Chem. Commun.*, **3**, 86–89.

[73] H. K. Ravi, M. Stach, T. A. Soares, T. Darbre, J.-L. Reymond, and M. Cascella, (2013), "Electrostatics and flexibility drive membrane recognition and early penetration by the antimicrobial peptide dendrimer bH1", *Chem. Commun.*, **49**, 8821–8823.

[74] M. Stach, T. N. Siriwardena, T. Köhler, C. van Delden, T. Darbre, and J.-L. Reymond, (2014), "Combining topology and sequence design for the discovery of potent antimicrobial peptide dendrimers against multidrug-

resistant *Pseudomonas aeruginosa*", *Angew. Chem., Int. Ed.*, **53**, 12827–12831.

[75] G. A. Eggimann, S. Buschor, T. Darbre, and J.-L. Reymond, (2013), "Convergent synthesis and cellular uptake of multivalent cell penetrating peptides derived from Tat, Antp, pVEC, TP10 and SAP", *Org. Biomol. Chem.*, **11**, 6717–6733.

[76] G. A. Eggimann, E. Blattes, S. Buschor, R. Biswas, S. M. Kammer, T. Darbre, and J.-L. Reymond, (2014), "Designed cell penetrating peptide dendrimers efficiently internalize cargo into cells", *Chem. Commun.*, **50**, 7254–7257.

[77] N. Martinho, H. Florindo, L. Silva, S. Brocchini, M. Zloh, and T. Barata, (2014), "Molecular modeling to study dendrimers for biomedical applications", *Molecules*, **19**, 20424–20467.

[78] M. Ballauff and C. N. Likos, (2004), "Dendrimers in solution: Insight from theory and simulation", *Angew. Chem., Int. Ed.*, **43**, 2998–3020.

[79] R. Bauer, V. Enkelmann, U. M. Wiesler, A. J. Berresheim, and K. Müllen, (2002), "Single-crystal structures of polyphenylene dendrimers", *Chem. - Eur. J.*, **8**, 3858–3864.

[80] G. Michaud, R. Visini, M. Bergmann, G. Salerno, R. Bosco, E. Gillon, B. Richichi, C. Nativi, A. Imberty, A. Stocker, T. Darbre, and J.-L. Reymond, (2016), "Overcoming antibiotic resistance in *Pseudomonas aeruginosa* biofilms using glycopeptide dendrimers", *Chem. Sci.*, **7**, 166–182.

[81] F. Vögtle, M. Ballauf, R. M. Crooks, B. I. Lemon, M. Möller, K. Müllen, D. Muscat, S. S. Sheiko, R. A. van Benthem, T. Weil, *et al.*, (2003), *Dendrimers III: Design, Dimension, Function*. Topics in Current Chemistry, Springer Berlin.

[82] S. V. Lyulin, A. A. Darinskii, and A. V. Lyulin, (2005), "Computer simulation of complexes of dendrimers with linear polyelectrolytes", *Macromolecules*, **38**, 3990–3998.

[83] Y. Liu, V. S. Bryantsev, M. S. Diallo, and W. A. Goddard III, (2009), "PAMAM dendrimers undergo pH responsive conformational changes without swelling", *J. Am. Chem. Soc.*, **131**, 2798–2799.

[84] J. Kłos and J.-U. Sommer, (2009), "Properties of dendrimers with flexible spacer-chains: a Monte Carlo study", *Macromolecules*, **42**, 4878–4886.

[85] G. M. Pavan, A. Barducci, L. Albertazzi, and M. Parrinello, (2013), "Combining metadynamics simulation and experiments to characterize dendrimers in solution", *Soft Matter*, **9**, 2593–2597.

[86] P.-G. de Gennes and H. Hervet, (1983), "Statistics of «starburst» polymers", *J. Phys., Lett.*, **44**, 351–360.

[87] R. L. Lescanec and M. Muthukumar, (1990), "Configurational characteristics and scaling behavior of starburst molecules: a computational study", *Macromolecules*, **23**, 2280–2288.

[88] P. K. Maiti, T. Cagin, S.-T. Lin, and W. A. Goddard, (2005), "Effect of solvent and pH on the structure of PAMAM dendrimers", *Macromolecules*, **38**, 979–991.

[89] B. Wu, W.-R. Chen, T. Egami, X. Li, Y. Liu, Y. Wang, C. Do, L. Porcar, K. Hong, L. Liu, *et al.*, (2012), "Molecular dynamics and neutron scattering study of the dependence of polyelectrolyte dendrimer conformation on counterion behavior", *J. Chem. Phys.*, **137**, 064902.

[90] S. H. Kim and M. H. Lamm, (2012), "Multiscale modeling for host-guest chemistry of dendrimers in solution", *Polymers*, **4**, 463–485.

[91] H. Lee and R. G. Larson, (2008), "Lipid bilayer curvature and pore formation induced by charged linear polymers and dendrimers: the effect of molecular shape", *J. Phys. Chem. B*, **112**, 12279–12285.

[92] B. Nandy and P. K. Maiti, (2010), "DNA compaction by a dendrimer", *J. Phys. Chem. B*, **115**, 217–230.

[93] W.-d. Tian and Y.-q. Ma, (2013), "Theoretical and computational studies of dendrimers as delivery vectors", *Chem. Soc. Rev.*, **42**, 705–727.

[94] L. Cavallo and F. Fraternali, (1998), "A molecular dynamics study of the first five generations of poly (propylene imine) dendrimers modified with N-tBoc-L-Phenylalanine", *Chem. - Eur. J.*, **4**, 927–934.

[95] B. P. Roberts, M. J. Scanlon, G. Y. Krippner, and D. K. Chalmers, (2009), "Molecular dynamics of poly(l-lysine) dendrimers with naphthalene disulfonate caps", *Macromolecules*, **42**, 2775–2783.

[96] B. P. Roberts, G. Y. Krippner, M. J. Scanlon, and D. K. Chalmers, (2009), "Molecular dynamics of variegated polyamide dendrimers", *Macromolecules*, **42**, 2784–2794.

[97] D. Moiani, M. Salvalaglio, C. Cavallotti, A. Bujacz, I. Redzynia, G. Bujacz, F. Dinon, P. Pengo, and G. Fassina, (2009), "Structural characterization of a protein a mimetic peptide dendrimer bound to human IgG", *J. Phys. Chem. B*, **113**, 16268–16275.

[98] S. Javor and J.-L. Reymond, (2009), "Molecular dynamics and docking studies of single site esterase peptide dendrimers", *J. Org. Chem.*, **74**, 3665–3674.

[99] D. M. Zuckerman, (2011), "Equilibrium sampling in biomolecular simulation", *Annu. Rev. Biophys.*, **40**, 41–62.

[100] L. C. S. Filipe, M. Machuqueiro, and A. M. Baptista, (2011), "Unfolding the conformational behavior of peptide dendrimers: Insights from molecular dynamics simulations", *J. Am. Chem. Soc.*, **133**, 5042–5052.

[101] S. Falkovich, D. Markelov, I. Neelov, and A. Darinskii, (2013), "Are structural properties of dendrimers sensitive to the symmetry of branching? Computer simulation of lysine dendrimers", *J. Chem. Phys.*, **139**, 064903.

[102] L. C. S. Filipe, M. Machuqueiro, T. Darbre, and A. M. Baptista, (2013), "Unraveling the conformational determinants of peptide dendrimers using molecular dynamics simulations", *Macromolecules*, **46**, 9427–9436.

[103] A. Hinchliffe, (1996), *Modelling molecular structures*. Wiley tutorial series in theoretical chemistry, Wiley.

[104] W. F. van Gunsteren and H. J. C. Berendsen, (1990), "Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry", *Angew. Chem., Int. Ed. Engl.*, **29**, 992–1023.

[105] C. Mura and C. E. McAnany, (2014), "An introduction to biomolecular simulations and docking", *Mol. Simul.*, **40**, 732–764.

[106] M. P. Allen and D. J. Tildesley, (1989), *Computer Simulation of Liquids*. Oxford Science Publ, Clarendon Press.

[107] D. Frenkel and B. Smit, (2001), *Understanding Molecular Simulation: From Algorithms to Applications*. Computational science series, Elsevier Science.

[108] A. R. Leach, (2001), *Molecular Modelling: Principles and Applications*. Prentice Hall.

[109] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, *et al.*, (2006), "Biomolecular modeling: goals, problems, perspectives", *Angew. Chem., Int. Ed.*, **45**, 4064–4092.

[110] J. W. Ponder and D. A. Case, (2003), "Force fields for protein simulations", *Adv. Protein Chem.*, **66**, 27–85.

[111] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, (1995), "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules", *J. Am. Chem. Soc.*, **117**, 5179–5197.

[112] J. Wang, P. Cieplak, and P. A. Kollman, (2000), "How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?", *J. Comput. Chem.*, **21**, 1049–1074.

[113] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, (1996), "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids", *J. Am. Chem. Soc.*, **118**, 11225–11236.

[114] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, (2004), "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6", *J. Comput. Chem.*, **25**, 1656–1676.

[115] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren, (2011), "Definition and testing of the GROMOS force-field versions 54A7 and 54B7", *Eur. Biophys. J.*, **40**, 843–856.

[116] M. M. Reif, P. H. Hünenberger, and C. Oostenbrink, (2012), "New interaction parameters for charged amino acid side chains in the GROMOS force field", *J. Chem. Theory Comput.*, **8**, 3705–3723.

[117] I. G. Tironi, R. Sperb, P. E. Smith, and W. F. van Gunsteren, (1995), "A generalized reaction field method for molecular dynamics simulations", *J. Chem. Phys.*, **102**, 5451–5459.

[118] T. Darden, D. York, and L. Pedersen, (1993), "Particle mesh Ewald: An $Nlog(N)$ method for Ewald sums in large systems", *J. Chem. Phys.*, **98**, 10089–10092.

[119] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, (1995), "A smooth particle mesh Ewald method", *J. Chem. Phys.*, **103**, 8577–8593.

[120] A. J. Ballard, S. Martiniani, J. D. Stevenson, S. Somani, and D. J. Wales, (2015), "Exploiting the potential energy landscape to sample free energy", *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, **5**, 273–289.

[121] M. Karplus and J. A. McCammon, (2002), "Molecular dynamics simulations of biomolecules", *Nat. Struct. Mol. Biol.*, **9**, 646–652.

[122] D. A. McQuarrie, (1975), *Statistical mechanics*. Harper's chemistry series, Harper & Row.

[123] R. W. Hockney, S. P. Goel, and J. W. Eastwood, (1974), "Quiet high-resolution computer models of a plasma", *J. Comput. Phys.*, **14**, 148–158.

[124] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, (2014), "Why the Monte Carlo method is so important today", *Wiley Interdiscip. Rev.: Comput. Stats.*, **6**, 386–392.

[125] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, (1953), "Equation of state calculations by fast computing machines", *J. Chem. Phys.*, **21**, 1087–1092.

[126] A. Baptista, (1998), *Theoretical Methods for the Simulation of Proteins at Constant pH*. PhD thesis, ITQB-UNL.

[127] M. Schaefer, M. Sommer, and M. Karplus, (1997), "pH-dependence of protein stability: Absolute electrostatic free energy differences between conformations", *J. Phys. Chem. B*, **101**, 1663–1683.

[128] A. M. Baptista, V. H. Teixeira, and C. M. Soares, (2002), "Constant-pH molecular dynamics using stochastic titration", *J. Chem. Phys.*, **117**, 4184–4200.

[129] D. Chandler, (1987), *Introduction to Modern Statistical Mechanics*. Oxford University Press.

[130] P. H. Hünenberger, (2005), "Thermostat algorithms for molecular dynamics simulations", in *Advanced Polymer Science*, vol. 173, 105–149, Springer.

[131] J. Wereszczynski and J. A. McCammon, (2012), "Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition", *Q. Rev. Biophys.*, **45**, 1–25.

[132] C. C. Moore, (2015), "Ergodic theorem, ergodic theory, and statistical mechanics", *Proc. Natl. Acad. Sci.*, **112**, 1907–1911.

[133] A. M. Baptista, P. J. Martel, and C. M. Soares, (1999), "Simulation of electron-proton coupling with a Monte Carlo method: Application to cytochrome c3 using continuum electrostatics", *Biophys. J.*, **76**, 2978–2998.

[134] G. M. Ullmann and E. Bombarda, (2013), "pK_a values and redox potentials of proteins. what do they mean?", *Biol. Chem.*, **394**, 611–619.

[135] E. Alexov, E. L. Mehler, N. Baker, A. M Baptista, Y. Huang, F. Milletti, J. E Nielsen, D. Farrell, T. Carstensen, M. H. M. Olsson, *et al.*, (2011), "Progress in the prediction of pK_a values in proteins", *Proteins: Struct., Funct., Bioinf.*, **79**, 3260–3275.

[136] A. V. Onufriev and E. Alexov, (2013), "Protonation and pK changes in protein-ligand binding", *Q. Rev. Biophys.*, **46**, 181–209.

[137] P. Ren, J. Chun, D. G. Thomas, M. J. Schnieders, M. Marucho, J. Zhang, and N. A. Baker, (2012), "Biomolecular electrostatics and solvation: a computational perspective", *Q. Rev. Biophys.*, **45**, 427–491.

[138] Y. Nozaki and C. Tanford, (1967), "Examination of titration behavior", *Methods Enzymol.*, **11**, 715–734.

[139] M. Machuqueiro and A. M. Baptista, (2011), "Is the prediction of pK_a values by constant-pH molecular dynamics being hindered by inherited problems?", *Proteins: Struct., Funct., Bioinf.*, **79**, 3437–3447.

[140] P. J. Martel, A. Baptista, and S. B. Petersen, (1996), "Protein electrostatics", vol. 2 of *Biotechnology Annual Review*, 315–372, Elsevier.

[141] B. Honig and A. Nicholls, (1995), "Classical electrostatics in biology and chemistry", *Science*, **268**, 1144–1149.

[142] A. Redondo and R. LeSar, (2004), "Modeling and simulation of biomaterials", *Annu. Rev. Mater. Res.*, **34**, 279–314.

[143] V. H. Teixeira, C. A. Cunha, M. Machuqueiro, A. S. F. Oliveira, B. L. Victor, C. M. Soares, and A. M. Baptista, (2005), "On the use of different dielectric constants for computing individual and pairwise terms in Poisson-Boltzmann studies of protein ionization equilibrium", *J. Phys. Chem. B*, **109**, 14691–14706.

[144] A. Warshel and J. Aqvist, (1991), "Electrostatic energy and macromolecular function", *Annu. Rev. Biophys. Biophys. Chem.*, **20**, 267–298.

[145] C. N. Schutz and A. Warshel, (2001), "What are the dielectric "constants" of proteins and how to validate electrostatic models?", *Proteins: Struct., Funct., Bioinf.*, **44**, 400–417.

[146] F. Fogolari, A. Brigo, and H. Molinari, (2002), "The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology", *J. Mol. Recognit.*, **15**, 377–392.

[147] D. Bashford and M. Karplus, (1990), "pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model", *Biochemistry*, **29**, 10219–10225.

[148] A. M. Baptista and C. M. Soares, (2001), "Some theoretical and computational aspects of the inclusion of proton isomerism in the protonation equilibrium of proteins", *J. Phys. Chem. B*, **105**, 293–309.

[149] M. S. Lee, F. R. Salsbury, and C. L. Brooks, (2004), "Constant-pH molecular dynamics using continuous titration coordinates", *Proteins: Struct., Funct., Bioinf.*, **56**, 738–752.

[150] J. Mongan, D. A. Case, and J. A. McCammon, (2004), "Constant pH molecular dynamics in generalized Born implicit solvent", *J. Comput. Chem.*, **25**, 2038–2048.

[151] S. Donnini, F. Tegeler, G. Groenhof, and H. Grubmüller, (2011), "Constant pH molecular dynamics in explicit solvent with λ-dynamics", *J. Chem. Theory Comput.*, **7**, 1962–1978.

[152] S. G. Itoh, A. Damjanović, and B. R. Brooks, (2011), "pH replica-exchange method based on discrete protonation states", *Proteins: Struct., Funct., Bioinf.*, **79**, 3420–3436.

[153] J. A. Wallace and J. K. Shen, (2011), "Continuous constant pH molecular dynamics in explicit solvent with pH-based replica exchange", *J. Chem. Theory Comput.*, **7**, 2617–2629.

[154] J. M. Swails, D. M. York, and A. E. Roitberg, (2014), "Constant pH replica exchange molecular dynamics in explicit solvent using discrete protonation states: Implementation, testing, and validation", *J. Chem. Theory Comput.*, **10**, 1341–1352.

[155] M. Machuqueiro and A. M. Baptista, (2006), "Constant-pH molecular dynamics with ionic strength effects: Protonation-conformation coupling in decalysine", *J. Phys. Chem. B*, **110**, 2927–2933.

[156] M. Machuqueiro and A. M. Baptista, (2008), "Acidic range titration of HEWL using a constant-pH molecular dynamics method", *Proteins: Struct., Funct., Bioinf.*, **72**, 289–298.

[157] H. A. F. Santos, D. Vila-Viçosa, V. H. Teixeira, A. M. Baptista, and M. Machuqueiro, (2015), "Constant-pH MD simulations of DMPA/DMPC lipid bilayers", *J. Chem. Theory Comput.*, **11**, 5973–5979.

[158] D. Vila-Viçosa, V. H. Teixeira, A. M. Baptista, and M. Machuqueiro, (2015), "Constant-pH MD simulations of an oleic acid bilayer", *J. Chem. Theory Comput.*, **11**, 2367–2376.

[159] D. Bashford and M. Karplus, (1991), "Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation", *J. Phys. Chem.*, **95**, 9556–9561.

[160] F. Zeng and S. C. Zimmerman, (1997), "Dendrimers in supramolecular chemistry: from molecular recognition to self-assembly", *Chem. Rev.*, **97**, 1681–1712.

[161] J. Kofoed and J.-L. Reymond, (2005), "Dendrimers as artificial enzymes", *Curr. Opin. Chem. Biol.*, **9**, 656–664.

[162] Y. Kim, F. Zeng, and S. C. Zimmerman, (1999), "Peptide dendrimers from natural amino acids", *Chem. - Eur. J.*, **5**, 2133–2138.

[163] M. J. Cloninger, (2002), "Biological applications of dendrimers", *Curr. Opin. Chem. Biol.*, **6**, 742–748.

[164] B. Romestand, J.-L. Rolland, A. Commeyras, G. Coussot, I. Desvignes, R. Pascal, and O. Vandenabeele-Trambouze, (2010), "Dendrigraft poly-L-lysine: A non-immunogenic synthetic carrier for antibody production", *Biomacromolecules*, **11**, 1169–1173.

[165] R. S. Navath, A. R. Menjoge, B. Wang, R. Romero, S. Kannan, and R. M. Kannan, (2010), "Amino acid-functionalized dendrimers with heterobifunctional chemoselective peripheral groups for drug delivery applications", *Biomacromolecules*, **11**, 1544–1563.

[166] N. Maillard, T. Darbre, and J.-L. Reymond, (2009), "Identification of catalytic peptide dendrimers by "off-bead" in silica high-throughput screening of combinatorial libraries", *J. Comb. Chem.*, **11**, 667–675.

[167] E. Kolomiets, E. M. Johansson, O. Renaudet, T. Darbre, and J.-L. Reymond, (2007), "Neoglycopeptide dendrimer libraries as a source of lectin binding ligands", *Org. Lett.*, **9**, 1465–1468.

[168] S. Javor, A. Natalello, S. M. Doglia, and J.-L. Reymond, (2008), "α-helix stabilization within a peptide dendrimer", *J. Am. Chem. Soc.*, **130**, 17248–17249.

[169] H. Lee and R. G. Larson, (2011), "Effects of PEGylation on the size and internal structure of dendrimers: Self-penetration of long PEG chains into the dendrimer core", *Macromolecules*, **44**, 2291–2298.

[170] K. Karatasos, (2008), "Self-organization in dendrimer polyelectrolytes", *Macromolecules*, **41**, 1025–1033.

[171] R. Blaak, S. Lehmann, and C. N. Likos, (2008), "Charge-induced conformational changes of dendrimers", *Macromolecules*, **41**, 4452–4458.

[172] Schrödinger, LLC, (2010), *The PyMOL Molecular Graphics System, Version1.7r1*.

[173] D. van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, (2005), "GROMACS: Fast, flexible, and free", *J. Comput. Chem.*, **26**, 1701–1718.

[174] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, (2008), "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation", *J. Chem. Theory Comput.*, **4**, 435–447.

[175] H. J. Berendsen, D. van der Spoel, and R. van Drunen, (1995), "GROMACS: A message-passing parallel molecular dynamics implementation", *Comput. Phys. Commun.*, **91**, 43–56.

[176] P. E. Smith and W. F. van Gunsteren, (1994), "Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K", *J. Chem. Phys.*, **100**, 3169–3174.

[177] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, (1984), "Molecular dynamics with coupling to an external bath", *J. Chem. Phys.*, **81**, 3684–3690.

[178] H. J. C. Berendsen, J. P. M. Postma, W. van Gunsteren, and J. Hermans, (1981), "Interaction models for water in relation to protein hydration", *Intermolecular Forces, Pullman B. Ed.*, **14**, 331–342.

[179] C. Cantor and P. R. Schimmel, (1980), *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*. Biophysical chemistry, W. H. Freeman.

[180] B. Efron and R. J. Tibshirani, (1994), *An Introduction to the Bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis.

[181] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, (1995), "Funnels, pathways, and the energy landscape of protein folding: A synthesis", *Proteins: Struct., Funct., Bioinf.*, **21**, 167–195.

[182] F. E. Cohen and M. J. E. Sternberg, (1980), "On the prediction of protein structure: The significance of the root-mean-square deviation", *J. Mol. Biol.*, **138**, 321–333.

[183] S. R. R. Campos and A. M. Baptista, (2009), "Conformational analysis in a multidimensional energy landscape: Study of an arginylglutamate repeat", *J. Phys. Chem. B*, **113**, 15989–16001.

[184] R. C. Pani and Y. G. Yingling, (2012), "Role of solvent and dendritic architecture on the redox core encapsulation", *J. Phys. Chem. A*, **116**, 7593–7599.

[185] H. Lee, J. S. Choi, and R. G. Larson, (2011), "Molecular dynamics studies of the size and internal structure of the PAMAM dendrimer grafted with arginine and histidine", *Macromolecules*, **44**, 8681–8686.

[186] H. Lee and R. G. Larson, (2009), "Molecular dynamics study of the structure and interparticle interactions of polyethylene glycol-conjugated PAMAM dendrimers", *J. Phys. Chem. B*, **113**, 13202–13207.

[187] N. W. Suek, , and M. H. Lamm, (2006), "Effect of terminal group modification on the solution properties of dendrimers: A molecular dynamics simulation study", *Macromolecules*, **39**, 4247–4255.

[188] P. K. Maiti, T. Cagin, G. Wang, and W. A. Goddard, (2004), "Structure of PAMAM dendrimers: Generations 1 through 11", *Macromolecules*, **37**, 6236–6254.

[189] B. W. Silverman, (1986), *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis.

[190] C. L. Brooks III, (2002), "Protein and peptide folding explored with molecular simulations", *Acc. Chem. Res.*, **35**, 447–454.

[191] G. G. Maisuradze, A. Liwo, and H. A. Scheraga, (2010), "Relation between free energy landscapes of proteins and dynamics", *J. Chem. Theory Comput.*, **6**, 583–595.

[192] M. Rubinstein and R. H. Colby, (2003), *Polymer Physics*. OUP Oxford.

[193] R. Scherrenberg, B. Coussens, P. van Vliet, G. Edouard, J. Brackman, E. de Brabander, and K. Mortensen, (1998), "The molecular characteristics of poly(propyleneimine) dendrimers as studied with small-angle neutron scattering, viscosimetry, and molecular dynamics", *Macromolecules*, **31**, 456–461.

[194] D. Bashford and K. Gerwert, (1992), "Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin", *J. Mol. Biol.*, **224**, 473–486.

[195] A. Luzar and D. Chandler, (1993), "Structure and hydrogen bond dynamics of water-dimethyl sulfoxide mixtures by computer simulations", *J. Chem. Phys.*, **98**, 8160–8173.

[196] A. Luzar and D. Chandler, (1996), "Hydrogen-bond kinetics in liquid water", *Nature*, **379**, 55–57.

[197] W. S. Hancock and J. E. Battersby, (1976), "A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid", *Anal. Biochem.*, **71**, 260–264.

[198] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, (2006), "Comparison of multiple amber force fields and development of improved protein backbone parameters", *Proteins: Struct., Funct., Bioinf.*, **65**, 712–725.

[199] B. Zagrovic, J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, S. Doniach, and V. S. Pande, (2005), "Unusual compactness of a polyproline type II structure", *Proc. Natl. Acad. Sci.*, **102**, 11698–11703.

[200] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, (1983), "Comparison of simple potential functions for simulating liquid water", *J. Chem. Phys.*, **79**, 926–935.

[201] G. Bussi, D. Donadio, and M. Parrinello, (2007), "Canonical sampling through velocity rescaling", *J. Chem. Phys.*, **126**, 014101.

[202] M. Parrinello and A. Rahman, (1981), "Polymorphic transitions in single crystals: A new molecular dynamics method", *J. Appl. Phys.*, **52**, 7182–7190.

[203] S. Nosé and M. L. Klein, (1983), "Constant pressure molecular dynamics for molecular systems", *Mol. Phys.*, **50**, 1055–1076.

[204] J. Rudnick and G. Gaspari, (1986), "The aspherity of random walks", *J. Phys. A: Math. Gen.*, **19**, L191–L193.

[205] I. Tanis and K. Karatasos, (2009), "Molecular dynamics simulations of polyamidoamine dendrimers and their complexes with linear poly(ethylene oxide) at different pH conditions: static properties and hydrogen bonding", *Phys. Chem. Chem. Phys.*, **11**, 10017–10028.

[206] V. Jain, V. Maingi, P. K. Maiti, and P. V. Bharatam, (2013), "Molecular dynamics simulations of PPI dendrimer-drug complexes", *Soft Matter*, **9**, 6482–6496.

[207] A. Ortega, D. Amorós, and J. G. de la Torre, (2011), "Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models", *Biophys. J.*, **101**, 892–898.

[208] G. D. R. Echenique, R. R. Schmidt, J. J. Freire, J. G. H. Cifre, and J. G. de la Torre, (2009), "A multiscale scheme for the simulation of conformational and solution properties of different dendrimer molecules", *J. Am. Chem. Soc.*, **131**, 8548–8556.

[209] I. Bodnár, A. S. Silva, R. W. Deitcher, N. E. Weisman, Y. H. Kim, and N. J. Wagner, (2000), "Structure and rheology of hyperbranched and dendritic polymers. i. modification and characterization of poly(propyleneimine) dendrimers with acetyl groups", *J. Polym. Sci., Part B: Polym. Phys.*, **38**, 857–873.

[210] I. B. Rietveld, , and D. Bedeaux, (2000), "Self-diffusion of poly(propylene imine) dendrimers in methanol", *Macromolecules*, **33**, 7912–7917.

[211] M. Z. Tien, A. G. Meyer, D. K. Sydykova, S. J. Spielman, and C. O. Wilke, 11 2013, "Maximum allowed solvent accessibilites of residues in proteins", *PLoS ONE*, **8**, e80635.

[212] M. An, J. M. Hutchison, S. R. Parkin, and J. E. DeRouchey, (2014), "Role of pH on the compaction energies and phase behavior of low generation PAMAM–DNA complexes", *Macromolecules*, **47**, 8768–8776.

[213] M. A. van Dongen, B. G. Orr, and M. M. Banaszak Holl, (2014), "Diffusion NMR study of generation-five PAMAM dendrimer materials", *J. Phys. Chem. B*, **118**, 7195–7202.

[214] K. Hong, Y. Liu, L. Porcar, D. Liu, C. Y. Gao, G. S. Smith, K. W. Herwig, S. Cai, X. Li, B. Wu, W.-R. Chen, and L. Liu, (2012), "Structural response of polyelectrolyte dendrimer towards molecular protonation: the inconsistency revealed by SANS and NMR", *J. Phys.: Condens. Matter*, **24**, 064116.

[215] I. Lee, B. D. Athey, A. W. Wetzel, W. Meixner, and J. R. Baker, (2002), "Structural molecular dynamics studies on polyamidoamine dendrimers for a therapeutic application: Effects of ph and generation", *Macromolecules*, **35**, 4510–4520.

[216] K. A. Sharp and B. Honig, (1990), "Electrostatic interactions in macromolecules: Theory and applications", *Annu. Rev. Biophys. Biophys. Chem.*, **19**, 301–332.

[217] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, (2006), "Modeling electrostatic effects in proteins", *Biochim. Biophys. Acta, Proteins Proteomics*, **1764**, 1647–1676.

[218] S. Kavyani, S. Amjad-Iranagh, and H. Modarress, (2014), "Aqueous poly(amidoamine) dendrimer G3 and G4 generations with several interior cores at pHs 5 and 7: A molecular dynamics simulation study", *J. Phys. Chem. B*, **118**, 3257–3266.

[219] G. M. Pavan, L. Albertazzi, and A. Danani, (2010), "Ability to adapt: Different generations of PAMAM dendrimers show different behaviors in binding siRNA", *J. Phys. Chem. B*, **114**, 2667–2675.

[220] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, (1997), "LINCS: A linear constraint solver for molecular simulations", *J. Comput. Chem.*, **18**, 1463–1472.

[221] S. Miyamoto and P. A. Kollman, (1992), "Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models", *J. Comput. Chem.*, **13**, 952–962.

[222] C. A. Carvalheda, S. R. R. Campos, M. Machuqueiro, and A. M. Baptista, (2013), "Structural effects of pH and deacylation on Surfactant Protein C in an organic solvent mixture: A constant-pH md study", *J. Chem. Inf. Model.*, **53**, 2979–2989.

[223] M. K. Gilson, K. A. Sharp, and B. H. Honig, (1988), "Calculating the electrostatic potential of molecules in solution: Method and error assessment", *J. Comput. Chem.*, **9**, 327–335.

[224] C. R. Martinez and B. L. Iverson, (2012), "Rethinking the term "π-stacking"", *Chem. Sci.*, **3**, 2191–2201.

[225] S. M. Malathy Sony and M. N. Ponnuswamy, (2006), "Nature of π-interactions in nitrogen-containing heterocyclic systems: A structural database analysis", *Cryst. Growth Des.*, **6**, 736–742.

[226] P. Chakrabarti and R. Bhattacharyya, (2007), "Geometry of nonbonded interactions involving planar groups in proteins", *Prog. Biophys. Mol. Biol.*, **95**, 83–137.

[227] J. Heyda, P. E. Mason, and P. Jungwirth, (2010), "Attractive interactions between side chains of histidine-histidine and histidine-arginine-based cationic dipeptides in water", *J. Phys. Chem. B*, **114**, 8744–8749.

[228] G. B. McGaughey, M. Gagné, and A. K. Rappé, (1998), "π-stacking interactions: Alive and well in proteins", *J. Biol. Chem.*, **273**, 15458–15463.

[229] T. Steiner and G. Koellner, (2001), "Hydrogen bonds with π-acceptors in proteins: frequencies and role in stabilizing local 3D structures", *J. Mol. Biol.*, **305**, 535–557.

[230] L. M. Salonen, M. Ellermann, and F. Diederich, (2011), "Aromatic rings in chemical and biological recognition: Energetics and structures", *Angew. Chem., Int. Ed.*, **50**, 4808–4842.

[231] A. Karshikoff and I. Jelesarov, (2008), "Salt bridges and conformational flexibility: Effect on protein stability", *Biotechnol. Biotechnol. Equip.*, **22**, 606–611.

[232] D. J. Barlow and J. M. Thornton, (1983), "Ion-pairs in proteins", *J. Mol. Biol.*, **168**, 867–885.

[233] C. A. Carvalheda, S. R. R. Campos, and A. M. Baptista, (2015), "The effect of membrane environment on Surfactant Protein C stability studied by constant-pH molecular dynamics", *J. Chem. Inf. Model.*, **55**, 2206–2217.

[234] R. H. Bartels, J. C. Beatty, and B. A. Barsky, (1995), *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*. Morgan Kaufmann Series in Computer Graphics and Geometric Modeling, Morgan Kaufmann.

[235] G. R. Grimsley, J. M. Scholtz, and C. N. Pace, (2009), "A summary of the measured pK values of the ionizable groups in folded proteins", *Protein Sci.*, **18**, 247–251.

[236] M. Lard, S. H. Kim, S. Lin, P. Bhattacharya, P. C. Ke, and M. H. Lamm, (2010), "Fluorescence resonance energy transfer between phenanthrene and PAMAM dendrimers", *Phys. Chem. Chem. Phys.*, **12**, 9285–9291.

[237] X. Shi, I. Lee, X. Chen, M. Shen, S. Xiao, M. Zhu, J. R. Baker, and S. H. Wang, (2010), "Influence of dendrimer surface charge on the bioactivity of 2-methoxyestradiol complexed with dendrimers", *Soft Matter*, **6**, 2539–2545.

[238] S. Riniker, A.-P. E. Kunz, and W. F. van Gunsteren, (2011), "On the calculation of the dielectric permittivity and relaxation of molecular models in the liquid phase", *J. Chem. Theory Comput.*, **7**, 1469–1475.

[239] P. Cieplak, W. D. Cornell, C. Bayly, and P. A. Kollman, (1995), "Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins", *J. Comput. Chem.*, **16**, 1357–1377.

[240] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, (1993), "General atomic and molecular electronic structure system", *J. Comput. Chem.*, **14**, 1347–1363.

[241] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, (1993), "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model", *J. Phys. Chem.*, **97**, 10269–10280.

[242] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, and M. Scharf, (1995), "The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies", *J. Comput. Chem.*, **16**, 273–284.

[243] U. Samanta, R. P. Bahadur, and P. Chakrabarti, (2002), "Quantifying the accessible surface area of protein residues in their local environment", *Protein Eng.*, **15**, 659–667.

[244] C. M. Topham and J. C. Smith, (2015), "Tri-peptide reference structures for the calculation of relative solvent accessible surface area in protein amino acid residues", *Comput. Biol. Chem.*, **54**, 33–43.

[245] V. H. Teixeira, D. Vila-Viçosa, A. M. Baptista, and M. Machuqueiro, (2014), "Protonation of DMPC in a bilayer environment using a linear response approximation", *J. Chem. Theory Comput.*, **10**, 2176–2184.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Apoio financeiro da FCT, bolsa nº SFRH/BD/76085/2011